

The
British Journal
of
Educational Psychology
(Incorporating the "Forum of Education")

EDITED BY

C. W. VALENTINE

(*Professor of Education in the University of Birmingham*)

WITH THE ASSISTANCE OF

R. L. ARCHER	LL. WYNN JONES
P. B. BALLARD	H. CRICHTON MILLER
CYRIL BURT	SIR T. PERCY NUNN
JAMES DREVER	H. BOMPAS SMITH
H. R. HAMLEY	C. E. SPEARMAN
SUSAN ISAACS	GODFREY THOMSON
HELEN M. WODEHOUSE	

AND THE FOLLOWING REPRESENTATIVES
OF OTHER COUNTRIES :

CHARLOTTE BÜHLER	PETER PETERSEN
ARNOLD GESELL	JEAN PIAGET
I. MEYERSON	L. M. TERMAN
E. L. THORNDIKE	

Volume IX
1939

Issued on behalf of the
BRITISH PSYCHOLOGICAL SOCIETY and the TRAINING COLLEGE
ASSOCIATION
by The British Journal of Educational Psychology, Ltd.
All Rights Reserved.

Reprinted with the permission of The British Journal of Educational Psychology Ltd.

JOHNSON REPRINT CORPORATION
111 Fifth Avenue, New York, N. Y. 10003

JOHNSON REPRINT COMPANY LIMITED
Berkeley Square House, London, W. 1

First reprinting, 1964, Johnson Reprint Corporation

British Journal of Educational Psychology.

INDEX TO VOLUME IX, 1939.

Articles :	PAGE
Classics, Performances in the Constituent Parts of a Higher School Certificate Examination in (James A. Petch)	174
Educational Abilities, The Relation of (Cyril Burt)	45
General Knowledge and Intelligence (Dorothea M. Inman)	72
Marking English Composition, The Efficiency of Different Methods of (B. M. D. Cast)	257
McDougall, William : An Appreciation (Cyril Burt)	1
Punishment in Schools (Kenneth D. Hopkins)	8
Scholastic Success at the Beginning and End of a Secondary School Career, The Bearing of General and Special Abilities Upon : Part I Part II (Mary Ormiston)	164 213
Scientific Interests of Senior School Children, The (R. Rallison)	117
Secondary Schools, A Study of Careers of Pupils who enter after their Second Attempt in the Admission Examination (D. M. Coombes)	145
Sex Differences among Indian Children in the Binet Simon Tests (V. V. Kamat)	251
Spelling Tests, Experiment with Non-Dictated (Stanley D. Nisbet)	29

	PAGE
Spens Report, Modern Psychology and The.. .. . (Olive A. Wheeler)	224
Teaching Profession, A Study of the Factors Useful in Choosing Candidates for the (J. Lawton)	131
Training College Students, Educational Abilities of (P. E. Vernon)	233
Critical Notices :	
Piaget, Jean : La Naissance de l'Intelligence chez l'Infant.. .. (M. M. Lewis)	92
Spearman, C. : Psychology Down the Ages (L. Wynn Jones)	88
Spens Report, The (P. B. Ballard)	196
Thomson, Godfrey H. : The Factorial Analysis of Human Ability (Cyril Burt)	188
Thurstone, L. L. : Primary Mental Abilities. (H. Eysenck)	270
Summaries of Researches Reported in Degree Theses	97 to 104, 201, and 276 to 278
Authors :	
Burt, Cyril : The Relation of Educational Abilities	45
Burt, Cyril : William McDougall : An Appreciation	1
Cast, B. M. D. : The Efficiency of Different Methods of Marking English Composition	257
Coombes, D. M. : A Study of Careers of Pupils who Enter Sec- ondary Schools after their Second Attempt in the Admission Examination	145
Hopkins, Kenneth D. : Punishment in Schools	8
Inman, Dorothea M. ; General Knowledge and Intelligence ..	72
Kamat, V. V. : Sex Differences Among Indian Children in the Binet Simon Tests	251

	PAGE
Lawton, J.: A Study of the Factors Useful in Choosing Candidates for the Teaching Profession	131
Nisbet, Stanley D.: Experiment with Non-Dictated Spelling Tests	29
Ormiston, Mary: The Bearing of General and Special Abilities upon Scholastic Success at the Beginning and End of a Secondary School Career: Part I	164
Part II	213
Petch, James A.: Performances in the Constituent Parts of a Higher School Certificate Examination in Classics	174
Rallison, R.: The Scientific Interests of Senior School Children ..	117
Vernon, P. E.: Educational Abilities of Training College Students	233
Wheeler, Olive A.: Modern Psychology and The Spens Report ..	224

WILLIAM McDUGALL : AN APPRECIATION.

By CYRIL BURT.

THE death of Professor McDougall, at the age of 67, has deprived the world of one of its best known psychologists and British psychology of one of its most stimulating leaders.

William McDougall was born in Lancashire on June 22nd, 1871. He was educated first at a private school near home and later at a German school at Weimar. At the early age of 15 he entered the University of Manchester. Four years later he went to St. John's College, Cambridge, and, after another four years, to St. Thomas's Hospital, London. His medical studies he regarded merely as a means to an end: even at this early age he had decided to devote himself, not to the practice of medicine, but to the exploration of the mental functions of the brain and nervous system. At the age of 27 he was elected to a Fellowship at St. John's College, but joined almost immediately the Cambridge Anthropological Expedition to Torres Straits, where he carried out a number of psychological tests on natives, and studied at first hand the organization of social life among the 'pagan tribes of Borneo.' On returning to Europe he worked for a while at Göttingen under G. E. Müller; and was then invited by Sully to undertake the teaching of experimental psychology at University College, London. In the development of experimental work in London he always retained the greatest interest: the first psychological laboratory was founded there by his initiative; and the earlier researches of Spearman, William Brown and other Londoners owed much to his keen interest and encouragement.

His first important publication (*Mind*, 1901) recorded a long series of ingenious experiments on colour vision, and sought to revise and reinstate Young's three-colour theory. From this he turned to a more ambitious study of "the physiological factors of the attention process" (*Mind*, 1902-3). This was part of an even larger problem that was occupying him at that time, namely, the desire to formulate a physiological basis for cognitive processes, in terms of an organized hierarchy of nervous arcs, through which the unknown nervous energy ("neurin" as he called it) flowed in a pattern of waves guided by the resistance of the nerve-junctions or synapses. His first book—the little *Primer of Physiological Psychology* (1905)—was a remarkably compact and lucid survey of mental activity described from this standpoint. Written under the inspiration of Sherrington's recent work on the "integrative action of

the nervous system," it sought to interpret Stout's cognitive psychology in terms of current neurological concepts and the dynamic psychology of James in terms of the Darwinian instincts, and to join the two into a coherent and consistent whole.

Meanwhile, in 1904, he had been appointed Wilde Reader in Mental Philosophy at Oxford—a post just vacated by Professor Stout. Here his further experiments on attention led him to the study of more abnormal states in which attention lapsed—hypnosis, suggestion, mental fatigue, and intoxication by alcohol and other drugs. Still seeking to lay bare the underlying machinery of mental life, he was, nevertheless, from the first convinced that mechanistic principles could never yield the final explanation. The brain and nervous system were the instruments of the mind, not the mind itself. By exhibiting the utmost that mechanism could do he believed we might best demonstrate that mechanism was not enough, and at last reach that "irreducible residuum" which implied for its explanation the co-operative agency of a purposive soul. *Body and Mind* (1911)—a work which he always looked back upon as his *magnum opus*—embodied these views in a large and systematic treatise, and sought to "establish by the methods of empirical science some foundation for the world-old belief that human personality is not wholly destroyed by death."

Nor did he consider it right to study mind solely in the abstract. Man he regarded as essentially a social individual; and both social psychology and individual psychology (which he conceived rather as a subdivision of social psychology than as an independent branch) claimed his attention from the outset. The practical applications of psychology—to education, to medicine, to social and industrial problems—always received his heartiest sympathy. "In psychology," he used to say, "there are so many unsolved problems that we may well devote ourselves to the useful ones first." Nevertheless, he remained a profound believer in the supreme importance of the experimental approach. Even the problem of the soul he thought should be solved by this method. "If ever I am seized by an incurable illness," he once announced to two startled students (who were not quite sure whether he spoke in earnest or in jest), "I have asked Sherrington to separate my cerebral hemispheres with his scalpel; and I hope I shall then be able to say whether I am still a single personality or two conscious streams." He was much impressed by Galton's work on mental testing, mental heredity, and correlation; and strongly supported the establishment of an Anthropometric Laboratory at Oxford. Though he himself was no statistician, he again and again predicted that statistics would play a large part in

psychology in the near future. "These mathematicians," he observed after a visit of Karl Pearson, "will soon be ousting us all from our jobs—physicists and psychologists alike."

He was the first in this country to realise the practical importance of the correlational work published by Spearman in American and German journals; and, largely as a result of his interest in the inheritability of mental differences and in the bearing of mental inheritance on eugenic and other social questions, he warmly encouraged his earlier students—William Brown, J. C. Flugel, May Smith, M. E. Bickersteth and myself—to investigate the practical possibilities of psychological testing. Among many fruitful suggestions for the measurement of innate intelligence, he himself devised an "alphabet test," a "spot pattern test," a portable tachistoscope, and, above all, the celebrated "dotting machine." His immediate aim was an "Anthropometric Survey of the British Isles," to be carried out through the schools and to review "both the physical efficiency and the mental efficiency" of the whole nation.

In 1912 he was elected a Fellow of the Royal Society. But at Oxford science and philosophy alike remained sceptical of any attempt to solve the problems of the mind by an experimental approach, particularly when introduced from Cambridge. My own tutor permitted me to attend McDougall's lectures; but added that I would "probably do better to spend any spare time on the river." "Nothing," I was assured, "will ever be discovered about the mind which is not already to be found in the writings of Aristotle." McDougall's lectures drew large crowds, and were the only lectures at which women predominated over men. "Is it," we were asked, "to gaze at that magnificent head, or to see two patients mesmerized upon the platform? Or do those deep notes mesmerize the audience, too?" His fine face and bearing, and almost Olympian manner, seemed immediately to set him apart as a Great Man—an effect which he secretly deplored. Yet, as Dr. Marett once observed, "he was indeed one of the great noblemen of science."

It is not easy to describe his own methods as a teacher. I well remember how we used to meet, almost in secrecy, like a pair of conspirators, in a large room in a deserted building at an hour when all healthy-minded undergraduates were "on the river." His air of lonely magnificence did little to dispel the sense of solitude. Chapter by chapter we worked through James's *Principles* and Ebbinghaus' *Grundzüge*. There was no systematic laboratory course in those days. Experimental technique, he thought, could best be acquired by acting as subject for actual researches. Among other items in my instruction it was proposed that I should study hypnotism at first hand by being hypnotised myself;

and, since the "Black Art" could not be practised on university premises, I used to bicycle to his house on Boar's Hill. There he became unexpectedly human and homely.

The experiments themselves were not without an educational interest. Their ultimate object was, first, to demonstrate that in the hypnotic state memory and other intellectual powers might be apparently improved and then to investigate whether the apparent improvement was due to the concentrated attention which such states seemed to entail or merely to the effects of suggestion. As a subject I myself was a failure. But the following year the experiments were resumed with Dr. May Smith as collaborator; and striking results were obtained, but (I believe) never published in any detail. In view of the popular notion of hypnotism as something occult, he feared they might easily be misinterpreted.

McDougall's experiments in hypnotism nearly cost him his post at Oxford; but the general theory of sleep, suggestion and hypnosis to which they led (*Brain*, 1908; *Encyclopædia Britannica*, s.v. 'Hypnotism'), and the attention that his work attracted to the whole subject, have had a manifest influence on the views of later writers on educational and medical psychology in this country.

Among the senior members of the University, his profoundest admirer next to Dr. Marett, was Maurice Keatinge, Reader in Education. Keatinge's own book on *Suggestion in Education* was, as he himself points out, largely inspired by McDougall's work. It was Keatinge, too, who supported the idea of introducing psychological tests into schools—in those days an almost unheard-of proposal—and persuaded the head masters of an elementary and a preparatory school ('Lynam's') to permit us to experiment on their pupils. McDougall was equally alive to the importance of applying similar methods to the study of children of pre-school age, and himself made an "experimental investigation into the colour sense of two infants." In the main, however, his own interest in mental testing lay rather in its social than in its educational applications.

From first to last, however, he took a special interest in the 'nature and technique of learning.' His theory of the learning process he later regarded as one of the chief 'novelties' that his work had contributed to psychology. In an important series of experiments, also carried out with Dr. May Smith, he endeavoured to show that memory is of two kinds—a mechanical memory, akin to habit, "belonging to the corpse," and a pure memory, inexplicable by mere association, "belonging to the soul within the corpse." "Our barbarian education," he wrote many years afterwards, "tacitly assuming that the pupil's mind is a mere

mechanism, places far too much reliance on memory of a mechanical kind: it was, perhaps, suited to the routine life of a simple Victorian age; it does not fit the pupil to play a civic part amid the complex social problems of the present century." And, again, "unintelligent learning involves no economy, no increase of efficiency, no achievement; intelligent learning involves achievement through insight, foresight and feeling. . . . Intelligent learning is the learning that must be introduced into the schools of the future."

All through his life he sought to stress the essential part that a science of psychology should play in all human and social problems. His chief aim was to fit it for this task. Psychology itself he regarded as a branch not of physiology but of biology; and he always insisted on the importance of a biological approach to the numerous problems of human nature. Although he welcomed and accepted Stout's contributions to cognitive or intellectual psychology, he felt that, in this country at any rate, the tendency had been to concentrate too exclusively on the intellectual side. Stout's own emphasis on conation to some extent corrected that bias; but Stout's interest lay in conation as it appears in cognitive life, not in conation studied in and for itself. The teaching of William James, with its scorn for intellectualism, greatly impressed McDougall; and he hoped to put dynamic psychology on a more systematic footing by describing the part played by instinct in all human behaviour. His book on human instincts and emotions was regarded by him, as the very title shows, primarily as an "introduction" to further studies on social psychology.

No work on psychology, and few serious works on science, can have passed through so many editions, or had have so wide an influence. It was primarily a comprehensive study of the nature and origins of motive. It combated the simple pleasure-and-pain theory of motive which the hedonistic philosophers had bequeathed both to education and to psychology. To the theorist it offered a suggestive biological account of the hereditary origin of these recurrent primitive motives, which seem so often at variance with the requirements of a civilised community; to the practical worker it supplied a most useful working classification for observing and recording those aspects of behaviour which are so often ignored in the classroom and yet become the centre of inquiry in the treatment of the neurotic and delinquent child. In its application to the problems of education, the purposive or hormic psychology, which it announced, was almost immediately accepted by such authoritative writers as Sir Percy Nunn and Maurice Keatinge; and through such channels as these McDougall's indirect influence on both the theory and the

practice of teaching must have been incalculable. His view provided precisely that psychological background which was needed to support the revolt against excessive intellectualism in education which was already gathering strength. He himself, as his sole direct contribution to educational writings amply shows (*Ped. Sem.*, XXXIII, pp. 203-314), was fully aware of the importance of "purposive experience" (as he called it in his title) in educational methods and aims.

McDougall was one of the first to welcome the work of Freud. With its insistence upon the importance of mental forces as distinct from mere ideas, and of unconscious processes as distinct from conscious states, the early teaching of Freud seemed closely in harmony with his own. It was, I imagine, partly because it was so easy to incorporate many of the simpler elements of psycho-analytic doctrine into the theory of mind already expounded by McDougall that psycho-analysis received an immediate sympathetic hearing among academic psychologists in this country.

McDougall's interest in abnormal psychology was still further increased when, owing to his medical qualifications, he was appointed a Major in the R.A.M.C., and with other psychologists was asked to undertake treatment of shell-shock casualties. More and more he became convinced that "perhaps the most important of all paths to a knowledge of the human soul is by way of psycho-pathology." Accordingly, when the war was over, he devoted the next few years to an attempt to "bridge the gap between academic psychology and the study of nervous and mental disorders."

In 1920 he accepted an invitation to occupy the Chair of Psychology at Harvard University. As the chair formerly held by William James, this ranked as the leading post in the psychological world of that day. A rapid series of books of a more popular and propagandist stamp than the earlier works on *Social Psychology* and *The Group Mind* (*National Welfare and National Decay*, 1921; *Ethics and some Modern World Problems*, 1924; *Janus, or the Conquest of War*, 1927; *Character and the Conduct of Life*, 1927; *World Chaos*, 1931) show that his interest in social problems was unabated. But in America he played the part rather of an indignant prophet calling the nations to repentance than the laboratory leader of a band of research-workers. At that time American psychology was greatly under the sway of the mechanistic teaching of the behaviourist school. McDougall, influenced largely by the views of James himself, remained convinced that "a materialistic mechanism was entirely inadequate to explain psychological phenomena." Accordingly, he now plunged into a sharp controversy with the mechanistic doctrines of the

psychologists around him; and his later books (*Modern Materialism and Emergent Evolution*, 1929; *Religion and the Sciences of Life*, 1934; *The Frontiers of Psychology*, 1934) are mainly absorbed with this problem. His experiments on the inheritance of acquired characteristics were animated by a kindred aim—to demonstrate that the evolution of the race, like the growth and progress of the individual, are the “embodiment of purpose.”

On leaving Harvard he arranged to spend half of each year in England, and rapidly began, so far as his increasing deafness would allow, to resume contact with his former fellow-workers and pupils. He came back saddened by private sorrows; but in the face of all his troubles he still preserved that air of noble fortitude which distinguished him throughout his life. In this country his strictures were chiefly directed against what he held to be the more extravagant aspects of the Freudian School. Returning to his old college in London, and speaking to a huge enthusiastic audience, he delivered his last set of public lectures in England. He chose for his subject, “Psychological Analysis and Social Psychology.” He still maintained that “Professor Freud has done more for the advancement of psychology than any student since Aristotle”; but he declared himself convinced (perhaps without quite realizing how far those doctrines had been modified) that many of the psycho-analytic doctrines—that of the *Œdipus complex* and its significance in early childhood, for example—were “misleading both to the teacher and to the doctor.”

Like James, he always took a profound interest in psychical research. At Duke University, during the closing years of his life, he did much to encourage the investigations of J. B. Rhine on extra-sensory perception. Here, again, by his courageous sympathy with possibilities that the scientific world had generally rejected, he laid himself open to a charge of being uncritical and unscientific himself; but it was, I fancy, rather an excess of the critical spirit than the lack of it that led him so frequently to challenge the established or orthodox standpoint. His last book of all, *The Riddle of Life* (1938), was completed after his partial recovery from a severe operation a year ago; and still dealt with the thesis that “mind is much more than mechanism, man much more than his body.” Here for the last time he proclaimed what he once called the ‘leading motive of his life’: “While I should prefer for myself a confident anticipation of total extinction at death to a belief that I must venture anew upon another life of whose nature and conditions I know nothing, I urgently desire, on impersonal grounds, to see the world-old belief in a future life established on a scientific foundation: the passing away of that belief would be calamitous for our civilisation.”

PUNISHMENT IN SCHOOLS.*

By KENNETH D. HOPKINS.

(*Education Department, Birmingham University.*)

- I.—Object.
- II.—Method (*Tables I and II*).
- III.—Summary of past and present work.
- IV.—Questionnaire.
- V.—Analysis of Questionnaire:
 - Section A.—*The supposed effects of different kinds of punishment. (Table III.)*
 - „ B.—*Should the home or the class be told of wrong-doing. (Table IV.)*
 - „ C.—*Punishments suggested by children for specific misdemeanours.*
 - „ D.—*Children's feelings when punished. (Table V.)*
 - „ E.—*The influence of the sex of the teacher.*
 - „ F.—*Children's ideas of the purpose of punishment.*
 - „ G and H.—*The personality of the teacher. (Table VI.)*
- VI.—Conclusions.

I.—OBJECT.

THE object of the enquiry here recorded was to determine the attitude of normal school children towards punishment and to find out in what respects their attitude changes with advancing age and education. By the term "normal school children" is meant those children who attend the ordinary primary, secondary, commercial and technical schools. It includes such problem children and difficult cases as are found in them and also those delinquents who, though on probation, are in attendance at the elementary schools. It excludes children who attend private schools and also those who, for feebleness of intellect, delinquency or any other cause, are dealt with in special schools or institutions.

II.—METHOD.

The principal method of enquiry was by the questionnaire given here.¹ This was supplemented by personal interviews and subsidiary questionnaires referred to or quoted where appropriate in the text.²

* An outline of a thesis approved as part qualification for the degree of M.A. in Education, Birmingham University.

¹ Pages 11-14.

² My most sincere thanks are due to the head masters and head mistresses of the schools which took part in this enquiry for their sympathy, co-operation and advice, and to my wife and Miss Oliver for their assistance in analysing the scripts.

Priority of place was given to the questionnaires rather than to the personal interviews because (a) it was thought that the children would feel more free to express themselves, (b) there would be less chance of influencing the children by suggestion or intuition, (c) in the interview, normal attitudes are apt to be overshadowed by abnormal idiosyncrasies, whereas the mass reactions obtained from the questionnaire emphasise normal group characteristics, which it was the object of this enquiry to determine.

The main questionnaire was the product of several trials applied in three schools and criticised by the head masters and staffs and, in the case of one secondary school, by some of the older pupils. It was found to be suitable for boys and girls between the ages of eight and seventeen years, but the story in Section A had to be modified for the girls by substituting the names "Mary, Jean and Elsie" instead of "Dick, Harry and Tom," as the girls were inclined to be excessively harsh in their judgments about boys. The girls' names were selected from among the commonest on the school register and, apart from twenty girls in the nine-year-old group from one school, who stated that Mary repeated her offence (cheating) "because she is always at it," there was no evidence that any personal significance was attached to any of the names. Reports from the schools showed that the children answered the questions with interest and enjoyment. Table I shows the types of schools selected for the experiment and Table II shows the age distribution of the children taking part in it.

TABLE I.
TYPES OF SCHOOLS USED.

<i>Schools</i>	<i>Remarks</i>
Boys' Secondary Girls' Secondary Mixed Secondary	Co-education might modify ideas on some aspects of school punishments.
Commercial Two Senior Elementary	Non-secondary school children, on the average of different environment, intelligence and social position.
Five Junior Elementary	Younger children from districts ranging from "good" to "very poor."
A London Junior Technical School	For comparison with Midland children.

TABLE II.
NUMBERS OF CHILDREN AT DIFFERENT AGES.

<i>Age in Years</i>	16+	15+	14+	13+	12+	11+	10+	9+	8+	<i>Total</i>
Boys	128	211	278	184	132	129	209	95	48	1414
Girls.....	87	147	180	216	146	134	174	96	34	1214
TOTAL										2628

III.—PAST AND PRESENT WORK.

The changed and changing attitude to punishment in schools is due to many influences. The chief, although perhaps the slowest and most gradual in effect, is the progress of human civilisation crystallised in, and partly directed by, the leaders of thought in ethics and social philosophy. These may be classed as "ideal views" that have gradually permeated all society and have affected ideas on punishment generally. Their tendency has been to diminish severity and lay emphasis on reformation. Of almost equal importance has been the modification of current practices due to increasing knowledge of the psychology of childhood and adolescence and the application of the principles enunciated by workers in these fields to practical pedagogy. Of recent years the Freudian doctrine of repression and its attendant mental ill-health has found ready acceptance in the flood of post-war sentimentalism and has resulted in such rapid changes that we are in danger of forgetting that "self-realisation does not necessarily demand freedom from restraint. In many cases it can be attained only by submitting to restraint."¹ Moreover, remembering how difficult it is for the adult to enter fully into the workings of the mind of a child, it is quite possible that the changes in ideas and practice which, to adults, have seemed so reasonable and so rich in promise of improvement may, to the immature mind of childhood, seem to be merely exhibitions of weakness or caprice.

These considerations justify the direct approach to those most concerned, the children themselves, for information regarding their thoughts about punishment, whether it has any restraining or reformatory influence and whether one type is more suitable than another. The problem has been approached by three different methods:

¹ J. ADAMS. *Evolution of Educational Theory*, p. 148.

- (a) Questionnaires or essays on selected topics.
- (b) Careful observation.
- (c) Personal interviews.¹

Among these the work of Dr. Cyril Burt is of outstanding interest, particularly in the emphasis it puts on the importance of considering the individual sinner rather than the misdemeanour for which he is being punished. However, it must not be forgotten that Burt's work was largely concerned with delinquent children whose special abnormalities required special remedial treatment, and when he writes "Delinquency I regard as nothing but an outstanding sample—dangerous perhaps and extreme but none the less typical—of common childish naughtiness,"² we are not justified in thinking, as some appear to think, that an outbreak of "common childish naughtiness" is a symptom of delinquency and should be treated as such. If we agree with the statement of Stanley Hall that "body growth is not symmetrical, but to some extent the parts, functions and organs grow in succession, so that the exact normal proportions of the body are temporarily lost, to be regained later on a new plan. The mind now grows in like manner. It is as if the various qualities of soul were developed successively . . ."³ we can readily believe that an outbreak of common childish naughtiness is, in most cases, due to a temporary, slightly excessive development of some mental function, which may, without detriment to mental health, be kept under control by judicious external measures until the development of the remainder of the mind is such that the child can once more take over its own control.

IV.—QUESTIONNAIRE.

This is not an examination. There are no right or wrong answers to the questions in this book. The object of this question paper is to find out what boys and girls of your age think about school punishments. I expect you are just as interested in this subject as I am. Will you answer the questions as carefully and thoughtfully as you can? In order that you may feel perfectly free to answer as frankly and honestly as possible, I shall not ask you to put your name on the paper and no one but me shall see your answers.

¹ *Pedagogical Seminary*: 1895, p. 235, "Punishment as seen by Children" (EARL BARNES); 1899, p. 159, "Home and School Punishments" (C. H. SEARS). *Journal of Education*: 1919, p. 91, "Children's Ideas on Punishment" (W. WOOD); 1921, p. 25, "Punishment" (J. L. PATON); 1927, p. 164, "Ethics of School Punishment" (S. ROWLAND).

² C. BURT. *The Young Delinquent*, p. viii.

³ STANLEY HALL. *Adolescence* Vol. II, p. 88.

Will you please give the following information?

Name of School :

Class : Age : years months.

Are you a boy or a girl?

A.—Once there were three boys. They were about as old as you^u are and their names were Dick, Harry and Tom. They all went to different schools and all did wrong and had to be punished sometimes. One day each of the boys was caught cheating by his teacher.

When the rest of the class went out to play, Dick's teacher called Dick to him and told him kindly how dishonest it was to cheat and pointed out that he would never do well at his lessons if he cheated. He then sent Dick out to play with the others.

Harry's teacher called Harry out and scolded him in front of the class and told the class what a mean and dishonest thing Harry had done.

Tom's teacher gave Tom a good caning.

A few days afterwards one of these three boys was caught cheating again.

- (1) Which boy do you think it was?
- (2) Why do you think so?
- (3) How do you think Dick felt after his punishment?
- (4) How do you think Harry felt after his punishment?
- (5) How do you think Tom felt after his punishment?
- (6) Whose teacher would you like best to teach you?
- (7) Why?
- (8) If you were the teacher and caught a boy cheating, what would you do?

B.—If your teacher caught you doing something very wrong indeed and said to you "I am going to tell either your parents or the rest of the class what you have been doing. Which of the two shall it be?" You *must* choose one of the two.

- (1) Which would you choose?
- (2) Why?

C.—Below is a list of six things for which boys and girls frequently get punished in school. Write down opposite to each one what punishment you would give for each. If you think no punishment should be given for some of them, write "none."

- | | |
|--------------------|---------------------|
| (1) Inattention : | (4) Forgetfulness : |
| (2) Disobedience : | (5) Telling lies : |
| (3) Bad work : | (6) Bullying : |

D.—(1) When you are punished do you feel

(a) Angry with yourself for having done wrong?

(b) Ashamed of yourself?

(2) Do you feel angry with your teacher for punishing you?

E.—(1) Which would you *dislike* most, to be punished by a man teacher or a lady teacher?

(2) Why would you *dislike* it most?

F.—(1) What is the good of punishment?

(2) Do you think teachers ought to punish you?

G.—(1) Does it make any difference to your feelings which teacher it is that punishes you?

(2) Why?

H.—This question is to find out what you think are the most important qualities a teacher should have. Below is a list of qualities arranged in pairs and you are asked to underline in each pair the one quality that you think is the more important.

For example, the first pair is: Strictness—Cleverness.

If you think strictness is more important than cleverness, underline *strictness*. If you think cleverness is the more important, underline *cleverness*. If you find it difficult to decide for any pair, be sure to underline one of them, even if you have to guess.

- | | |
|-------------------------------|-------------------------------|
| (1) Strictness—Cleverness. | (16) Good at sports— |
| (2) Cleverness—Manners. | Sense of humour. |
| (3) Manners—Dignity. | (17) Personal appearance— |
| (4) Dignity—Enthusiasm. | Sympathy. |
| (5) Enthusiasm—Fairness. | (18) Strictness—Dignity. |
| (6) Fairness—Sense of humour. | (19) Cleverness—Enthusiasm. |
| (7) Sense of humour— | (20) Manners—Fairness. |
| Sympathy. | (21) Dignity—Sense of humour. |
| (8) Sympathy—Good at sports. | (22) Enthusiasm—Sympathy. |
| (9) Good at sports— | (23) Fairness—Good at sports. |
| Personal appearance. | (24) Sense of humour— |
| (10) Manners—Strictness. | Personal appearance. |
| (11) Dignity—Cleverness. | (25) Enthusiasm—Strictness. |
| (12) Enthusiasm—Manners. | (26) Cleverness—Fairness. |
| (13) Fairness—Dignity. | (27) Manners—Sense of humour. |
| (14) Sense of humour— | (28) Dignity—Sympathy. |
| Enthusiasm. | (29) Good at sports— |
| (15) Sympathy—Fairness. | Enthusiasm. |

- | | |
|--|--|
| (30) Personal appearance—
Fairness. | (38) Good at sports—Manners. |
| (31) Strictness—Fairness. | (39) Dignity—
Personal appearance. |
| (32) Sense of humour—
Cleverness. | (40) Sympathy—Strictness. |
| (33) Manners—Sympathy. | (41) Cleverness—Good at sports. |
| (34) Good at sports—Dignity. | (42) Personal appearance—
Manners. |
| (35) Enthusiasm—
Personal appearance. | (43) Strictness—Good at sports. |
| (36) Strictness—
Sense of humour. | (44) Personal appearance—
Cleverness. |
| (37) Cleverness—Sympathy. | (45) Strictness—
Personal appearance. |

V.—ANALYSIS OF QUESTIONNAIRE.

Section A.—The supposed effects of different kinds of punishment.

Three types of punishment are presented for consideration: (i) Dick's—explanation, (ii) Harry's—public disgrace, (iii) Tom's—caning. The intention was to find out which of the three types of correction was thought to be most effective in preventing a recurrence of the fault and suggested a proper moral attitude. A rough measure of the estimated effectiveness of the punishments at each age can be obtained by subtracting the percentage of reported failures from 100. Table III shows the analysis of the most striking aspects of the answers to this section for the "explanation" method. As may be expected, the method appears to fail more frequently with younger children than with older ones, but there was only one reason given for this failure and that was "weakness." This was expressed in A2 in such phrases as "He only talked to him," "It was not a proper punishment," "She thought that if she had been let off once she would be again." One boy (age 10-1-) replied, "Dick cheated again because he wanted to go in front of the class for punishment," and another (age 10-1-), "Dick wanted to show his teacher he was not a sis" (i.e. a weakling). Several boys of twelve and fourteen and a few girls thought the teacher was "soft." The weakness of the method was shown to fall under three headings, (a) weakness of the punishment itself, (b) weakness on the part of the teacher, (c) failure to recognize that the culprit could endure a harder punishment.

The *Explanation Method* can be regarded as reasonably effective only above the age of 12-1-, for below that age more than half the children report that it failed to prevent a recurrence of the offence. The comparatively low figures of failure at the higher ages suggest the suitability of this method of correction for the older boys and girls, and the close

agreement with the figures for "indifference" at these ages suggests that the failure is due rather to moral backwardness on the part of the child than to weakness inherent in the method. But for twelve years old and under, some other explanation must be found, for the table shows that the punishment is increasingly thought to be ineffective, and yet not wholly because of indifference on the part of the children, e.g., for the 9+ boys, 61 per cent reported that the method was ineffective but only 20 per cent gave "indifference" as the reason; feelings of shame, sorrow and resolutions to reform were the generally reported results of the teacher's explanation of the fault.

TABLE III.

COMPARING FAILURE OF "EXPLANATION" METHOD AND ACCOMPANYING EMOTIONS.

Question	Boys	Percentage in each age group								
		16+	15+	14+	13+	12+	11+	10+	9+	8+
A2	Failure due to weakness	28	23	27	40	47	59	60	61	50
A3	Affect: Indifference . . .	24	19	28	30	42	36	27	20	25
	Shame and Sorrow	58	55	47	44	42	41	39	66	65
	Ideal	14	19	16	7	5	2	10	—	—
	GIRLS									
A2	Failure due to weakness	14	19	16	37	47	55	69	66	47
A3	Affect: Indifference . . .	10	12	14	15	33	16	35	14	6
	Shame and Sorrow	56	44	46	49	44	70	47	72	79
	Ideal	12	24	19	17	6	4	1	2	—

It appears that the child repeats his offence not because he does not care but because the punishment has failed to help him. There is some support for this view in the answers to A6 (which teacher would you like to teach you?) Although about 70 per cent of the children of all ages chose Dick's or Mary's teacher (the ones that explained), at fifteen years old 42 per cent of the boys and 47 per cent of the girls made this choice because the teacher "understood" and at ten years old 53 per cent of the boys and 49 per cent of the girls chose this teacher because he or she "gave no punishment" or was "kind." Considering the

inducement that there was to choose either Dick's or Mary's teacher and have an easy time, the numbers choosing the more severe teachers was relatively high, particularly on the grounds of "efficiency." In answer to A8 (what punishment would *you* give?) few of the younger ones contented themselves with mere explanation; the cane immediately, or explanation accompanied by caning or threats of the cane, were the most popular choices, and although some of the excessive severity must be attributed to childish barbarity, there was a strong suggestion that, for young children, if the method of explanation is employed, it must be supported by something more stimulating. Further evidence on this point was given in answer to F1 (what is the good of punishment?), where many of the younger ones stated that punishment is good "because it helps you to remember."

Certain answers to A3 were classed as "ideal" (see Table III). These denoted a spirit of co-operation between teacher and child together with a recognition of the fault; "Dick felt he had done wrong but, as his teacher was so decent, he must play the game also," "Mary felt her teacher still trusted her" were answers of this type given by a boy and a girl of 15+ respectively.

The punishment of *Public Disgrace* was thought effective by all ages, with the possible exceptions of eight and sixteen years. Answers to Section B suggest that children at these ages do not attach as much value to the opinions of their class-mates as do the children in the other age groups. This punishment and also "caning" were reported to fail at these two ages most frequently owing to the arousal of the instinct of self-assertion.

Corporal punishment was thought most effective by pupils below the age of 12+. It produced very mixed feelings, of which anger and revengefulness were among the commonest but, whereas these feelings were so pronounced among the elder children as to cause a repetition of the offence, among the younger they were acknowledged to be momentary and slight. For example, at 15+ the punishment was declared a failure by 51 per cent of the boys, 22 per cent giving revenge as the reason, while at the age of 10+ only 19 per cent declared failure, of which 2 per cent gave revenge as the reason. From the age of 14+ downwards an increasing number of boys and girls appeared to consider that a caning was simply a just retribution for the misdeed, wiping it out and leaving the sinner free to repeat the offence or not, as he pleased. In choosing the teacher preferred (in answer to A6 and 7, approximately 30 per cent of the children chose the severer teachers, particularly the one who used the cane, the deduction being clearly made in many cases that the

teacher who could punish efficiently could also teach efficiently and the class would make progress.

Though there was abundant evidence to show that corporal punishment should be avoided so far as possible and cautiously but effectively administered when to refrain would be a sign of weakness, there was no evidence that it evokes that feeling of physical and spiritual degradation that some of the more enthusiastic opponents of the practice assert. In one of the supplementary questionnaires about 300 children were asked to choose which punishment they would prefer out of the following : (1) two strokes with the cane, (2) five hundred lines to be done at home, (3) one half-day's detention in school. Of the boys 90 per cent, and of the girls 37 per cent, chose the cane, mostly because it was quickly over ; 10 per cent of the boys and 60 per cent of the girls chose lines and 3 per cent of the girls chose detention. Comparison of the choices made in the main and subsidiary questionnaires suggests that personal convenience (or inconvenience) plays a substantial part in a child's attitude towards punishment. In the main questionnaire the choice was between three punishments that took place during school hours and did not impose extra restraint on the children during their free time. The cane is quickly over, the public scolding took place during lesson time and the kindly lecture obviously detained the child for only a few minutes ; the lecture, which caused the least personal inconvenience, was the most popular choice. In the smaller experiment the choice was between the momentary pain of a sharp caning, the monotony of writing five hundred lines and the restraint of the detention. The popular choice was again for the punishment that entailed least inconvenience ; the boys chose the cane ; the girls, who would probably have to spend much of their evenings at home in any case, chose the lines. One very significant similarity appeared in the answers to both questionnaires—no girls above the age of fifteen years chose the cane nor, except in a few cases for bullying (Section C), did they recommend caning as a punishment. This is in conformity with the widely held belief that adolescent girls have a far greater reverence for their bodies than have boys. Consequently corporal punishment is likely to and, as the records show, does arouse feelings of anger to such an extent as to make it an unsuitable punishment for adolescent girls.

Section B.—Should the home or the class be told of a wrong-doing?

The purpose of this section was to find out the relative importance to the child of the opinions of parents and of class-mates or social equals. Answers taken to show that home opinion was the more important were

- (1) Ashamed for parents to know;
- (2) Parents would be sympathetic;
- (3) Telling the class would be no punishment, i.e., the opinion of equals did not matter.

Answers showing class opinion was of the greater importance were

- (1) Ashamed for children to know;
- (2) Class would be sympathetic;
- (3) Parents would keep the fault secret.

Approximately 70 per cent of the answers in each age group were included under these six headings. A positive or negative balance in favour of the importance of class opinion was found by subtraction. Table IV shows the results analysed on the scheme indicated above.

TABLE IV.
CHILDREN'S VALUES OF HOME AND CLASS OPINIONS.

Boys	Age	Percentage choosing at each age								
		16+	15+	14+	13+	12+	11+	10+	9+	8+
Home opinion important.....		45	34	33	23	26	23	21	35	33
Class opinion important.....		22	49	38	40	37	33	26	22	17
Balance in favour of opinion of class.....		-23	+15	+5	+17	+13	+10	+5	-13	-16
GIRLS										
Home opinion important.....		56	45	45	25	36	17	16	22	24
Class opinion important.....		33	42	44	49	43	42	44	30	18
Balance in favour of opinion of class.....		-23	-3	-1	+24	+7	+25	+18	+8	-6

The table shows that the importance attached to class opinion compared with home opinion increases from eight to thirteen years and then begins to decline. Some points of particular interest are noted below :

- (1) The small value attached to class opinion at 8-9 years is to be expected from observations of the instability of social groups at that age. The small value attached to class opinion and the increased value of home opinion at sixteen years may be accounted for under two heads :

- (a) the child is approaching adulthood and home conditions are conditions of social equality ;
 - (b) Sixteen years is the leaving age for secondary schools and therefore school society has become unstable.
- (2) The sudden drop in the balance in favour of class opinion for girls of 12+ and boys of 14+ is remarkable. The figure for class importance is still high, indicating that class opinion is still valuable but there is a sudden increase in the value attached to home opinion, particularly under the heading of " parents sympathetic." It is suggested that, at these ages, girls and boys are at a difficult period in their lives and the physical changes marking the onset of puberty not only evoke an exceptionally sympathetic attitude in the parents but make the child particularly aware of its need for sympathy.
- (3) The development of the girls is about one to two years in advance of the boys. There is a hint of this given in Table III, in the ages at which " ideal " relations between children and teacher are first recorded. Stronger confirmation is found in the analysis of Section D (personal feelings when punished).
- (4) In Section A, public disgrace appeared to be a most effective punishment and to evoke most intense feelings of shame and loss of self-respect at the ages 12-14 years, when the opinion of the class is shown to be of great importance.

Section C.—Punishments suggested by children for specific misdemeanours.

The punishments prescribed by the children were, in nearly all cases, the standard school punishments. Caning as a general remedy for all faults was more frequent among the younger than the older children. *Bullying* was accounted the worst fault and even girls of fifteen and sixteen united with boys in prescribing corporal punishment. The marked exceptions occurred among twelve-year-old boys, several of whom expressed the view that the bully should not be punished because the victim " should have fought back." This age group also contained two boys who, in Section B, would prefer to have the class told of their faults because they could fight the class. These instances, as well as the play of twelve-year-old boys, bear testimony to the strength of the instinct of pugnacity at that age. There was no indication in any age group that bullying was a moral offence. *Lying* appeared to be the next most serious fault and that too was most frequently punished by caning, except that about 10 per cent of the younger children prescribed writing " I must tell the truth " or some similar sentence several hundred times.

Among the older ones, particularly the girls, the moral aspect of lying was more and more generally recognized. 30 per cent of the boys and 70 per cent of the girls, aged fifteen and sixteen, suggested a confidential talk with the liar because "caning is not the proper punishment" and "to find out why he (or she) was telling lies."

Forgetfulness was seldom punished because "no one can help forgetting." When punishment was given it was always mild unless the forgetfulness was attributed to negligence in learning lessons, in which case detention or the cane was suggested.

Inattention was punished by a smack, or the cane, "not too hard," or by learning after school the work missed. One-third of the children prescribed no punishment either because the teacher ought to have made the lesson more interesting, or, less frequently, because the child would later receive the natural punishment in loss of marks. At all ages the punishments were roughly graded in severity and it was possible to arrange the offences in descending order of magnitude thus—

- (1) Bullying.
- (2) Telling lies.
- (3) Disobedience.
- (4) Bad work.
- (5) Inattention.
- (6) Forgetfulness.

Further points of interest that appeared during the analysis of this section are given below.

(1) If a caning is given it must be given properly and the teacher who cannot produce (in two or three shrewd blows) the proper amount of physical pain earns the contempt of the class.

(2) To be sent to the head master was accounted a particularly severe punishment by the pupils at the commercial and technical schools because it was feared that it would adversely affect their chances of being recommended for employment on leaving school.

(3) To be deprived of small privileges or pleasures was regarded as a severe punishment by the children of nine to eleven years in the schools in the poorest parts of the town. This suggests that the severity of this form of punishment depends on the circumstances and environment of the child.

(4) "Shouting at" was much dreaded. There was evidence that it is employed by many parents in the poorer homes and by a number of teachers. Remembering that, in the child's world, the adult is a symbol of stability and security, to see one apparently on the verge of losing all self-control must be, to the child, a terrifying spectacle.

Section D.—Feelings when punished.

About 70-80 per cent of the children of all ages and both sexes reported feeling angry with themselves either frequently or sometimes; 80-90 per cent reported that they felt ashamed and 25-35 per cent reported a feeling of anger against the teacher. That the thwarting of a wish is accompanied by a feeling of anger is a familiar idea which finds confirmation in the experimental work of H. L. Philp.¹ It has previously been suggested that childish naughtiness is a manifestation of instinctive urges temporarily out of control² and anger against the teacher, which, when reported at all, was frequently qualified as momentary, is here suggested to be the feeling-tone aroused by the thwarting of the progress towards the instinctive or immediate goal—an assumption supported by the fact that the higher percentage reporting this feeling was among the younger children, in whom instinctive forces are less under volitional control. The choice of a teacher who can help the pupils to "get on," the desire for an understanding teacher and the value put upon social teachers', or parents' opinions noticed in the answers to the previous sections and also noted by W. F. Brook in *The High School Teacher from the Pupils' Point of View*³ suggest that the main wish of the normal child (whether explicitly stated or not) is to make progress at school and become a useful member of society. It is suggested that anger against the self is the feeling-tone aroused by the momentary frustration of progress towards this end by the eruption of an instinctive urge. If these assumptions are correct, a comparison of the answers to D1 and D2 give a rough idea of the relative importance to the child of the instinctive and voluntary aims and suggest that, in the case of the normal child, the self is organised in the will to do right and that wrong-doing is impulsive.

Piaget⁴ shows the close connection between the development of the idea of justice and the progress of the child towards the co-operative stage of social relationship. Care was taken to avoid suggesting the idea of fairness or justice in any part of the questionnaire (except in Section H) but, in this section particularly, there was the opportunity for adding to their answers the words "if I deserve it." In compiling Table V the answers to Section D were not the only source of information. If a child, in any part of the paper, used any word or expression showing that the idea of punishment was associated with the idea of justice, credit was given for it.

¹ *British Journal of Psychology*. Monograph Suppt. XXI, 1936.

² This essay, p. 11.

³ *Pedagogical Seminary* 1905, pp. 239 *et seq.*

⁴ *The Moral Judgment of the Child*. Ch. III.

TABLE V.

CHILDREN ASSOCIATING PUNISHMENT WITH JUSTICE.

Sex	Percentage of children at each age								
	16+	15+	14+	13+	12+	11+	10+	9+	8+
Boys	50	44	33	31	30	18	14	5	—
Girls	50	43	41	47	36	26	19	13	6

The table shows a steady increase in the percentages mentioning justice with increasing age, the girls approximately a year ahead of the boys.

Section E.—The influence of the sex of the teacher.

Apart from the eight and nine-year-old groups, where the majority of boys and girls stated that they disliked being punished by a man more than by a woman because the punishment was usually more severe, the differences in opinion were so distinct that the sexes have to be taken separately. From nine to twelve years the opinion of the boys was fairly equally divided between men and women teachers but from thirteen years upwards punishment by a woman was more and more disliked. Up to twelve years the main reason dictating the choice was "severity," the more severe being the more disliked, but from thirteen years onwards this decreased in importance and reasons based on sex differences were increasingly given. Four distinct sex attitudes were noticed :

(1) Sex inferiority, generally plainly stated in some such expression as "women are inferior to men." This feeling, though small on the whole, reached its maximum at the ages of 13+ and 14+.

(2) Failure of women to understand the motives and feelings of boys, generally expressed in the form "women don't understand boys," or "men understand boys better." This feeling of lack of sympathetic understanding steadily increased with age.

(3) Dislike of women, given by the younger boys only.

(4) Chivalry, which increased with increasing age, was expressed in such terms as "it is unfair to play up with a woman," "women feel punishing boys more." Its first appearance was at eleven years old and its growth was rapid to the maximum at sixteen years. There was a most unexpected drop to half its expected value at the age of 14+. Separate analysis of the boys of the mixed secondary school, where

daily association with women teachers might be expected to modify their views, showed a similar drop in the chivalric attitude at 14+. The reason for this may be, as has been suggested for previous examples of anomalous behaviour in fourteen-year-old boys, those pubertal changes that, at their onset, may incline the boy towards a masterful rather than a protective attitude to women.

Girls from eight to thirteen years more frequently objected to being punished by a man than a woman, the objections reaching their maximum in the 12+ group, which has previously come under particular notice.¹ At fourteen, fifteen and sixteen the objections were about half and half, the sixteen-year-old group showing quite a high percentage of declared indifference. Among the objections based on sex reasons, that of lack of understanding was markedly lower at all ages than in the case of boys. At all ages a few, sometimes less than 1 per cent, stated that they "liked to be punished by a man," the number of cases rising suddenly to 10 per cent at the age of 16+. At this age only there were 15 per cent who baldly stated that they disliked men. In addition there were increases in the numbers reporting humiliation or embarrassment when punished by men. These rather mixed feelings in girls of sixteen may perhaps be the effect of the attainment of what is, in more primitive societies, marriageable age.

The general effect of the answers to this section is to raise a doubt as to whether the most beneficial results are obtained by continuing the co-educational system beyond the age of 11+, when children leave the junior elementary schools.

Section F.—Children's ideas of the purpose of punishment.

The answers to F1 (what is the good of punishment?) are summarised below :

- (1) About 30-40 per cent of the boys and 40-50 per cent of the girls regarded punishment as simply retributive. Answers of this type were "it pays you out," "it learns you."
- (2) The idea of reformation showed a progressive increase with age among the boys but kept a fairly constant and lower level among the girls.
- (3) Education as a "good" of punishment, noted particularly by A. C. Ewing² as applicable to the young, was expressed in statements such as "it teaches you right from wrong," "it teaches you to know wrong," "it teaches you what is right."

¹ This essay, p. 19.

² A. C. EWING: *The Morality of Punishment*, pp. 114 and 115.

- (4) Comparatively few of either sex expressed the true deterrent view that punishment is to stop others from wrong-doing.
- (5) About 10 per cent of the children stated that the good of punishment was to enforce discipline. This generally was on account of other unruly members of the class, indicated by the substitution of "them" or "the children" for the more personal "you" in their answers. This suggests that children realise that laws are framed for the protection of society and that it is justifiable to enforce them.
- (6) Very few indeed suggested that the purpose of punishment was to induce repentance.

A few children of all ages stated that punishment was "no good," but on the whole the answers to F1 clearly showed that children believe there is a purpose in punishment and that it is not solely the result of caprice; in fact, in several cases it was stated that the good of punishment was conditional upon its not being given to vent the teacher's temper.

The answers to F2 (ought the teacher to punish?) suggested that normal children of all ages accept the teacher's authority and accord to him (or her) the right to punish. The records suggest that failure to conform to this behaviour is not always a serious sign of abnormality in young children, but in adolescent children it should be regarded with misgiving.

Sections G and H.—The personality of the teacher.

About one-third of the boys and one-quarter of the girls gave negative replies to G1 (does it make any difference which teacher punishes you?). The older children mostly gave as their reason "the punishment is the same whoever gives it," suggesting that the fact that they deserved punishment was what really mattered. The more usual reason among the younger ones was "all teachers are alike," suggesting that their relationship with teachers was not fully co-operative. Accompanying the affirmative replies to G1, the principle reason, up to the age of 12+, was that some teachers are more harsh than others. Among the adolescents, affection for a teacher was increasingly given with advancing age. Several times a teacher who could make the punishment fit the crime was commended but teachers whose punishments made the children "feel small" were disliked. At all ages, but particularly with boys of thirteen and fourteen and girls of twelve, there were some children who showed concern about the teacher's regard for them. "The teacher does not respect you any more" or "some teachers don't like you after they have punished you" are samples of their answers. A few children

appeared to hold the view that offences were against the teacher and not against school society or moral laws and should be punished by the offended teacher and no one else.

In Section H, the qualities that seemed valuable in a teacher and, in the opinion of the pupils, tended to induce affection and respect were arranged in order of importance by the method of paired comparisons. Table VI shows the ranking of these qualities in order of importance by adolescent and pre-adolescent boys and girls and endorses the conclusion of W. F. Brook, "It is the common virtues and the more fundamental qualifications in teachers that appeal to high school children most. Personal appearance, neatness, politeness, culture, etc., however important for students in lower grades, are supplanted at the high school stage by more fundamental things."¹

TABLE VI.
TEACHERS' QUALITIES RANKED BY CHILDREN.

	Boys		Girls	
	<i>Adolescent</i>	<i>Pre-adolescent</i>	<i>Adolescent</i>	<i>Pre-adolescent</i>
1	Fairness	Manners	Fairness	Manners
2	Cleverness	Cleverness	Manners	Cleverness
3	Enthusiasm	Strictness	Sympathy	Fairness
4	Manners	Good at sport	Cleverness	Sense of humour
5	Sense of humour	Sense of humour	Sense of humour	Sympathy
6	Sympathy	Sympathy	Enthusiasm	Good at sport
7	Appearance	Fairness	Appearance	Appearance
8	Good at sport	Appearance	Good at sport	Enthusiasm
9	Strictness	Enthusiasm	Dignity	Strictness
10	Dignity	Dignity	Strictness	Dignity

The possession of all these virtues is not of itself sufficient to establish a full relationship of sympathetic co-operation between teacher and pupils. Occasionally, in answer to A6 (whose teacher would you like best?), children named some particular teacher at their own school and gave as their reason some apparently trivial property of the teacher

¹ *Pedagogical Seminary*, 1905: "The High-School Teacher from the Pupils' Point of View," p. 287.

concerned, such as "he lets us do fretwork," "he is good at stick-printing." It seemed as though the teacher, by his ability to do something that the child thought valuable and important, had bridged the gap between the adult and childish planes of social development and had made a spontaneous, probably unwitting, descent to the child's own level. I have even known the ability to referee a hockey match by whistling shrilly through his teeth, when the official whistle had been lost, earn for a master a greater measure of respect and esteem from his pupils than had ever been accorded to him for his academic qualifications. Stanley Hall quotes that Wickern, the founder of the Rauhe Haus, near Hamburg, for poor and exposed children, considered that no teacher who could not enter into the plays of childhood was fit to come in contact with them.¹

VI.—CONCLUSIONS.

(1) Lack of discipline and restraint by parents and teachers constitute a hardship in the case of the normally developing child. If it is admitted that even adults, at times, fail in the moral power to control their impulses, it is important to consider whether, by withdrawing compulsion and putting the onus of good behaviour on the child, a responsibility is being imposed that the powers of childhood are not sufficiently developed to support. With children under twelve, at least, argument and reasoning appear to be insufficient to produce a lasting effect.

(2) The social development of a child is progressive and anything tending to retard such development should be avoided. Public disgrace, which separates the offender from his group, is likely to induce the feeling that there are two external coercive forces, the teacher and the class, opposing him and that he can co-operate with neither. If deterrence is the object in administering a public scolding, the recognition of this aspect of punishment by the children seems to be too slight to guarantee sufficient success to balance its disadvantages.

(3) Corporal punishment should be administered only rarely to adolescents. The head master, who, by virtue of his office, is aloof from the society of all but the oldest pupils, is the most suitable person to inflict it. For while the class teacher's endeavour is towards social co-operation the head master remains as the symbol of coercive rule for those who still require it. On the other hand, with younger children, the sharp corrective and stimulating effect of an occasional stroke with the cane is thought to be of greater benefit than the most sympathetic explanations that rely on powers not yet developed.

¹ STANLEY HALL: *Adolescence*, Vol. I, p. 395.

(4) To avoid creating a feeling of injustice, punishment should conform, as far as possible, to the socially approved procedure. Young children particularly recommended as "proper punishment" first a clear warning, second the warning carried into effect if the fault is repeated.

(5) Feelings of shame, anger, revenge and sorrow are usually of short duration and are, to some extent, the natural consequences of the frustration of an impulse. Unless additional causes of aggravation are present they do not constitute a greater menace to social and moral development than would the unchecked offence.

(6) Prior to adolescence the sex of the teacher is immaterial but with older children it would be an advantage for boys and girls to be educated separately and to be controlled by teachers of their own sex.

(7) Children on the whole recognize and accept the authority of a teacher but their affection for and the influence of the teacher depend on the possession of certain qualities that the children feel to be important. The opportunity for the full exercise of these qualities comes when the gap between the adult and childish planes is successfully bridged.

(8) The attitude of children towards punishment changes with age and education. Changes are particularly noticeable in respect of regard for justice and social disapproval, the teacher's sex and the rating of the teacher's qualities. Ages requiring special attention are, for boys, fourteen and fifteen and, for girls, twelve to thirteen years.

Résumé.

DE LA PUNITION DANS LES ÉCOLES

Cet article décrit une enquête qui avait pour but de constater l'attitude des élèves anglais normaux envers la punition, et de découvrir de quelle façon cette attitude se transforme à mesure que l'âge augmente. L'on demanda à plus de deux mille élèves des deux sexes et de rang social varié, entre l'âge de 8 et 16 ans, de fournir des réponses anonymes à un questionnaire. Des résultats l'on tira les conclusions suivantes.

- (1) Les élèves reconnaissent en général l'autorité du maître et son droit de punir, mais l'efficacité de la punition dépend jusqu'à un certain degré du cas que fait l'élève de son maître.
- (2) Les explications et les appels à la raison sont efficaces avec les adolescents mais les enfants plus jeunes ont besoin en outre d'une punition pour produire des résultats efficaces.
- (3) Aux-dessus de l'âge de douze ans les enfants tendent à montrer quelque ressentiment contre la punition par un professeur de l'autre sexe.
- (4) Les sentiments de colère, de vengeance et de honte ne sont que légers et transitoires pourvu que la punition soit convenable et méritée.
- (5) Le désir qu'une punition soit impartiale ou juste augmente avec l'âge et avec le développement de la conscience sociale.

ZUSAMMENFASSUNG.

STRAFEN IN DER SCHULE.

Beschreibt eine Untersuchung, um die Stellungnahme normaler englischer Schulkinder der Strafe gegenüber zu ermitteln und herauszubekommen, auf welche Art und Weise die Stellungnahme sich mit zunehmendem Alter ändert. Über 2,000 Kinder beiden Geschlechts und verschiedenen sozialen Rangs im Alter von 8-16 Jahren wurden gebeten, anonym auf einen Fragebogen zu antworten. Danach ergab sich folgendes :

- (1) Die Macht des Lehrers und sein Recht zu bestrafen wird allgemein anerkannt, aber die Wirksamkeit der Strafe hängt in gewissem Masse von der Achtung des Kindes vor dem Lehrer ab.
- (2) Auseinandersetzungen und Appelle an die Vernunft sind bei Adoleszenten wirksam, aber die jüngeren Kinder brauchen die Strafe noch dazu, um wirksame Resultate hervorzurufen.
- (3) Über 12 Jahre sind die Kinder geneigt, einem Lehrer des anderen Geschlechts die Strafe übelzunehmen.
- (4) Gefühle des Zornes, der Rache, und der Scham sind nur gering und vorübergehend, solange die Strafe passend und verdient ist.
- (5) Das Verlangen, dass eine Strafe anständig und gerecht sein soll, nimmt mit dem Alter und mit dem Wachsen gesellschaftlichen Bewusstseins zu.

NON-DICTATED SPELLING TESTS.

BY STANLEY D. NISBET.

(*From Moray House Demonstration School, Edinburgh.*)

- I.—*Object of the investigation.*
- II.—*The types of test to be examined.*
- III.—*Previous opinions and investigations.*
- IV.—*Justification for the investigation.*
- V.—*Method.*
- VI.—*Comparison I—Number of errors.*
- VII.—*Comparison II—Number of "similar" responses.*
- VIII.—*Comparison III—Correlation.*
- IX.—*Comparison IV—Extent of influence of certain disadvantageous features in the "Wrongly-Spelt Word" and "Skeleton Word" types of test.*
- X.—*Deciding between "Skeleton Word" and "Wrongly-Spelt Word" types of test—points favouring final selection of "Skeleton Word" type as best.*
- XI.—*Conclusions.*

I.—OBJECT OF THE INVESTIGATION.

THE object of this experiment was to estimate the validity of three types of non-dictated spelling test, using dictation as a criterion assumed to be valid, and thus to show which of the three is the best substitute for a dictated test in cases where the latter is impracticable.

II.—THE TYPES OF TEST TO BE EXAMINED.

A prominent feature of our educational system is the examination for children of about eleven or twelve years of age, held at the end of the primary school course, by means of which entrants to post-primary schools are classified. In any of these examinations one of the most important sections is always an English test, and it is with the spelling items in these tests that this investigation has to deal.

The test questions in spelling present a special difficulty. It seems to be generally agreed that the best way of testing spelling is the dictation, preferably in a context, of words which are to be written down. Such a method, however, not only involves too many practical difficulties when large numbers of pupils are sitting the examination at the same time, often

in different schools, but does not allow of complete uniformity of administration owing to the differences of pronunciation and manner of presentation between those who administer the tests. Various devices have therefore been adopted to circumvent the problem by means of *non-dictated* ("self-administering" is the term used by American investigators) spelling tests included among the other test questions in the English paper.

There are several types of non-dictated spelling test. Of the eight types of spelling test mentioned by T. G. Foran in his *Psychology and Teaching of Spelling*,¹ five do not depend on dictation, and other forms than those mentioned by him have appeared from time to time. I have selected what I consider to be the only three types of non-dictated test which could reasonably be offered as a possible substitute for a dictated test. These are :

(a) *The "Multiple Choice" test* (hereafter referred to as *M.C.*). In this test several spellings of each word are given, and the correct one has to be underlined. For example : *I can play the (paino, piono, pieano, piano, peano, piaino).*

This form of test has been widely used both in this country and in America. The number of different spellings given of each word varies from two in some tests to six in others. Obviously, the smaller the range of choice, the larger the influence of the chance factor, and although this influence can be reduced by a modified method of scoring, the efficacy of the test is not so complete, while the more complex method of scoring is open to practical objections. In my tests I used the six-word form, and thus reduced the influence of chance as much as was practically possible.

(b) *The "Skeleton Word" test* (hereafter referred to as *Skel.*) In this test only so many letters of the test word are given as to ensure recognition, i.e., realization of what the word is. For example : *They are not friends : they often (qu r l) with one another.* The pupil is to write the word in full on the line at the end of the sentence. It will be seen that spaces are left where letters are omitted, but the exact number of letters required to fill these spaces is not indicated.

Unlike the *M.C.* test, the *Skel.* test is not at all widely used. It has several forms : sometimes only the first and last letters are given ; sometimes dots are put in to represent the missing letters. Believing that the words were in the former case too difficult to recognise, and in the latter case too easy to spell, I compromised by giving several letters of the word, with completely blank spaces (always a little larger than what was required to fill them) where letters were omitted.

¹ Page 179. (Publ. by the Catholic Education Press, Washington, D.C.)

(c) *The "Wrongly-Spelt Word" test* (hereafter referred to as *W.Sp.*) In this test the pupil has to write correctly a word which he is told is wrongly spelt. For example: *The Romans came to (conker) Britain.*

The correction of wrongly spelt words forms the basis of many varieties of spelling test. Of these I chose that form in which one misspelt word is given in a sentence. This is the form which T. G. Foran finds "has greater reliability and validity than any other test,"¹ and gives results near enough to the other forms of test in which errors have to be corrected for it to stand as a representative of that particular type.

As stated at the outset, the object of this experiment was to compare each of these types with a criterion assumed to be valid, viz., the method of dictation.

III.—PREVIOUS OPINIONS AND INVESTIGATIONS.

Many complaints have from time to time been made about these tests, especially by primary school teachers who prepare their pupils for the examinations referred to above. For example, some teachers, remembering the confusion they themselves experience when correcting many papers from bad spellers, say the M.C. test is too difficult, as confusing the pupil. Others say it is too easy: the pupil may be able to recognise a correct spelling, but not write it out for himself; they agree with Boyd that "a child cannot spell a word correctly until he can write it down correctly spelled."² As for the Skel. test, it is objected that it depends too much on intelligence: a child who could spell the word may not be able to discover what the word is. And in the *W.Sp.* test, the possible distorting effect of seeing words wrongly spelt naturally evokes some misgiving in those acquainted with the workings of the child mind.

One of the most relevant of previous investigations on this subject is by W. S. Guiler.³ He has compared M.C., *W.Sp.*, and dictation, and gives some important tables of results. His subjects range from the 7th grade (i.e., age about 14 years) to college students. However, what prevents me from accepting his conclusions is his adoption, as the criterion of validity, "the potency with which the different methods discover words that students cannot spell." That is to say, the type of test which gives the greatest number of errors is the best. But may such a test not show most errors because it contains elements which mislead students who can spell the word correctly under normal circumstances? If this is so, what is being measured includes ability to resist wrong suggestions

¹ *Loc. cit.*

² *Measuring Devices in Composition, Spelling and Arithmetic.* (Harrap and Co.)

³ *Journal of Educational Research*, XX, Oct., 1929, pp. 181-9.

(probably this is included in Guiler's "mastery" of a word), an ability which does not seem to me to be by any means essential in a good speller. A good speller is usually one who writes down the correct spelling automatically, and his ability, or lack of it, to resist wrong suggestions is really a side-issue, though a closely related one.

In an article entitled "A Comparison of Two-Response and Dictated Recall Types of Spelling Test"¹ D. P. Phillips compares dictated and M.C. tests. I agree with him that the criterion to be used is dictation ("Dictated-Recall"), but the value of his results is impaired by the fact that in the M.C. test which he used only two spellings were given for each word, one right and one wrong. Such a test is hardly discriminative enough, for in most cases there are several common mis-spellings of a word. Phillips' subjects, too, were all college students, whereas this investigation concerns children of about eleven.

W. W. Cook,² in a detailed large-scale experiment, compared a number of non-dictated ("self-administering") tests with a dictated test as a criterion, and proved conclusively that "recall" tests (i.e., those in which the word has to be written down correctly) are all more valid than "recognition" tests (of which I am using the M.C. as the representative in the present experiment).

T. G. Foran,³ besides giving very important results concerning the various forms of dictated test, investigated the M.C. test and two forms of W.Sp. test. His table of correlations shows, as did Cook's results, that the W.Sp. ("recall") tests give results nearer to those of dictated tests than does the M.C. ("recognition") test. Foran's results are specially noteworthy for the present investigation because his subjects, in the part of his experiment which included non-dictated tests, were children in the 6th, 7th, and 8th grades (i.e., ages about 13, 14, and 15).

R. Pintner, H. D. Rinsland, and Joseph Zubin, in an article entitled "The Evaluation of Self-Administering Spelling Tests,"⁴ compare two forms of M.C. test with dictation.

A similar comparison of M.C. ("Recognition") and dictation ("Recall") was made by Miss Evelyn M. Sturdyvin.⁵ I refer to her results later, but in this case too the subjects were college students.

¹ *Journal of Educational Research*, XXIII, 1931, pp. 17-24.

² "Measurement of General Spelling Ability, Involving Controlled Comparisons between Techniques," University of Iowa.—*Studies in Education*, Vol. 6, No. 6, 1932.

³ *The Psychology and Teaching of Spelling*, pp. 180-183.

⁴ (*American*) *Journal of Educational Psychology*, XX, 1929, pp. 107-111.

⁵ (*American*) *Journal of Educational Psychology*, May, 1937. Vol. XXVIII, No. 5, pp. 394 ff.

IV.—JUSTIFICATION FOR THE INVESTIGATION.

The investigation seems to be justified on the following grounds :

- (1) The complaints of teachers and educational authorities must be met.
- (2) In previous investigations little (if any) attention has been paid to the possibilities of a test of the " Skeleton-Word " type.
- (3) Little work seems to have been done on the subject in this country.
- (4) Most previous experiments appear to have been made on subjects older than those we have under consideration.

V.—METHOD.

(1) Eighty pupils were taken for the experiment. The distribution of their ages was as follows :

Lowest age :	10 years.
Highest age :	13 years 7 months.
Mean age :	11 years 7 months.
Standard deviation :	11·3 months.

(A frequency distribution table of the ages is given in Appendix I.)

(2) One hundred words, selected from Boyd's Spelling List,¹ were dictated (with three rest periods) to these 80 pupils. A context was given with each word, though only the word had to be written down. (The full list of words is given in Appendix II.)

The distribution of scores in this dictation was as follows :

Lowest mark :	23 out of 100.
Highest mark :	93 out of 100.
Mean mark :	60·7 out of 100.
Standard deviation :	19·65.

(A frequency distribution table of the scores is given in Appendix III.)

It was decided that this distribution showed that the words used were of just the right difficulty, and also that they had spread the pupils sufficiently, to produce good results.

(3) On the results of this dictation, the pupils were divided into four groups—Group I, Group II, Group III, Group IV—of equal spelling ability, with 20 pupils in each group. The groups were equated with regard to all details—total marks, range, and standard deviation.

(4) The same 100 words, from the results of a previous dictation in another school, had already been divided into four lists—List A, List B, List C, List D—of approximately equal difficulty, with 25 words in each list.

¹ Published by Harrap and Co.

(5) Twelve non-dictated tests were composed. Of each of the three types (M.C., Skel., W.Sp.) there were four tests, one for each of the four lists of words. Thus each test had 25 words in it.

The details of test-construction were as follows :

M.C. tests.—The alternative spellings, which were given along with the correctly spelt word, consisted in each case of the five most common mis-spellings of the word in question.

Skel. tests.—As far as possible, the only letters of a word which were given were those common to almost all the attempts to spell it, i.e., those about which there seemed to be little or no doubt.

W.Sp. tests.—The wrong spelling was made as unlike the true spelling as possible, in order to minimise "disturbing influence" (discussed later).

(6) The same words were given to the same pupils four days (two school days) after the dictation by means of the non-dictated tests mentioned above, and also by a second dictation, according to the following plan :

<i>M.C. test :</i>	List A	given to	Group I.
	List B	" "	Group II.
	List C	" "	Group III.
	List D	" "	Group IV.
<i>Skel. test :</i>	List A	" "	Group IV.
	List B	" "	Group I.
	List C	" "	Group II.
	List D	" "	Group III.
<i>W.Sp. test :</i>	List A	" "	Group III.
	List B	" "	Group IV.
	List C	" "	Group I.
	List D	" "	Group II.
<i>2nd Dict. :</i>	List A	" "	Group II.
	List B	" "	Group III.
	List C	" "	Group IV.
	List D	" "	Group I.

It will be seen that each type of test involved all the words and all the pupils, although each word was given to each pupil only once. The following possible distorting factors were thus eliminated :

- (a) Any list of words being especially suitable or unsuitable for any type of test ;
- (b) Any group of pupils being especially suited or unsuited to any type of test ;

- (c) The effect of memory. Because of the large number of the words, the youth of the pupils, and the interval between the day of the dictation and the day of the tests, the possibility of a pupil's remembering on the latter day how he spelt a word on the former day must have been fairly remote (but see below). But if he had been presented with the same word, albeit in different types of test, several times on the latter day, the effect of his memory would probably have been considerable.

The criterion used was the original dictation, consisting of 8,000 responses in all (80 pupils, 100 words). Thus, whatever group a pupil was in, or whatever list was used in any of the tests he had to do, a comparison with his performance, in the same words, in the original dictation was always possible.

The second dictation, which, like each of the non-dictated tests, consisted of 2,000 responses in all, was not, of course, used as the criterion, but it was nevertheless an integral part of the experiment. It was given under exactly the same conditions as the non-dictated tests with the object of finding *the maximum reliability which could be expected from any spelling test under these conditions, i.e., of finding how much variation from the original dictation could be accounted for merely by the natural instability of the spelling of children of that age.* Therefore when any comparison of a test with the original dictation was being made, the results of the second dictation were always taken into consideration. For example, if the M.C. test was under consideration, the process would be somewhat as follows :

- (1) Comparison of M.C. and original dictation gives a certain result.
- (2) Comparison of 2nd dictation with original dictation gives a certain result.
- (3) The latter comparison gives the highest degree of correspondence which could be expected.
- (4) Does the former result come near enough to it ?

To put it in another way, it is not so much the original dictation which is the criterion as the comparison between it and the second dictation which is the criterion. Thus we cancelled out not only the effect of the instability of the children's spelling referred to above, but also the effect of any change from the conditions of the original dictation (e.g., those due to practice-effect).

VI.—COMPARISON I—NUMBER OF ERRORS.

The first comparison was by the total number of errors made in the tests of each type.

It will be remembered (p. 34) that for each of the non-dictated tests and for the second dictation a combination of four groups of pupils and four lists of words was used. Any such combination, it will be seen, consisted of 2,000 responses in all.

We require now to get a figure, representing the original dictation, with which to compare the error totals in the non-dictated tests. In the original dictation (8,000 responses in all) there were 3,142 errors in all. Remembering that we are using parallel lists and parallel groups we may assume that in the original dictation the number of errors in each 2,000-response combination of four groups of pupils and four lists of words is in the region of 785 (quarter of 3,142). With this figure, therefore, we can compare the other results.

TABLE I.

TOTAL NUMBER OF ERRORS MADE IN EACH TYPE OF TEST.

Original Dictation....	c.785 errors in (each group of) 2,000 responses.
M.C. test	584 errors in 2,000 responses.
Skel. test.....	799 „ „ „
W.Sp. test	828 „ „ „
2nd dictation	709 „ „ „

From the above table it appears that (a) of the three non-dictated tests, the Skel. test shows up best (i.e., nearest to the original dictation) ; and (b) the M.C. and W.Sp. tests give too few and too many errors respectively. When we take into account the changed conditions apparently indicated by the decrease in errors in the second dictation, the Skel. test still shows up best.

The accuracy of result (b) is borne out by all the investigators mentioned before, whose tables of results also show the smallest number of errors when an M.C. test is given, and the largest number when a W.Sp. test is given, and by a previous investigation of my own, along similar lines, made in another school. This gives the lie to the criticism, already referred to, that the M.C. test is confusing and difficult, and vindicates the other view, that recognising a correct spelling is easier than writing it down.

VII.—COMPARISON II—NUMBER OF "SIMILAR" RESPONSES.

In the non-dictated tests and the second dictation I divided, for the purposes of this comparison, the responses into "similar" and "dissimilar" responses.

I have called a response "similar" if it gave the same result as the corresponding (i.e., same pupil, same word) response in the original dictation, that is, when the response and its corresponding response in the original dictation were either both correct or both wrong.

In the same way, I have called a response "dissimilar" when it gave a different result from its corresponding response in the original dictation, that is, when either it was wrong and its corresponding response was correct, or vice versa.

TABLE II.

"SIMILAR" AND "DISSIMILAR" RESPONSES IN EACH TYPE OF TEST.

<i>Type of Test.</i>	<i>Total No. of Responses.</i>	<i>No. of "similar" Responses.</i>	<i>No. of "dissimilar" Responses.</i>	<i>Percentage of "similar" Responses.</i>
M.C.....	2000	1521	479	76.1
Skel.	2000	1634	366	81.7
W.Sp.	2000	1666	334	83.3
2nd Dict.	2000	1729	271	86.5

This comparison is probably the best of them all, for it provides exactly what is wanted—the number, or percentage, of responses giving the same result in the test in question as in the criterion, the original dictation. Again the W.Sp. and Skel. tests are best, and are very close to each other, but the W.Sp. test has a slight advantage. The largest percentage of "similar" responses which could be expected is seen to be 86.5, and both the Skel. and the W.Sp. tests come so close to this that their validity is obvious.

VIII.—COMPARISON III—CORRELATION.

In the following table of correlation coefficients (product-moment) it should be noted that the scores in each test are correlated not with the total scores in the original dictation but only with the scores in the "corresponding questions" in the original dictation. (It has been pointed out before that each response in the non-dictated tests or the second dictation has a corresponding response in the original dictation, i.e., one involving the same pupil and the same word.)

TABLE III.
CORRELATION BETWEEN THE TESTS AND DICTATION.

Correlation between M.C. and original dictation	= .87 ± .019
" " Skel. " " "	= .90 ± .014
" " W.Sp. " " "	= .91 ± .013
" " 2nd dict. " " "	= .92 ± .011

The highest correlation which could be expected is, from the table, .92. All the tests come sufficiently near this to prove their validity; even .87 for the M.C. test would be, on commonly accepted standards, quite high enough to justify substitution of M.C. for dictation as a test. The coefficients for the W.Sp. and Skel. tests come very near to the maximum of .92, and their superiority is thus again shown.

The partial correlation coefficients, with age held constant, were found to differ from the above only in the third decimal place.

The superiority of "recall" over "recognition" tests is agreed to by other investigators (although the Skel. form of test is not mentioned by them), and we need not hesitate to accept the above results as fairly accurate.

Comparing other investigators' results in more detail, we note that between M.C. and dictation Phillips¹ finds a correlation of .69 ± .02, and Foran² a correlation of .756. The low figure in the former case is probably due to the fact that a 2-word form of the M.C. test was used; chance would thus be a disturbing factor. Even in Foran's investigation a 4-word form was used, and here chance may have had more effect than in the 6-word form of the M.C. in the present experiment. Miss Sturdyvin³ found a correlation of .884 ± .008, also using a 4-word form of the test, but all her subjects were college students, from whom a higher correlation was only to be expected, since their spelling is less unstable than that of children of 10-13.

IX.—COMPARISON IV—EXTENT OF INFLUENCE OF CERTAIN DISADVANTAGEOUS FEATURES IN THE "SKELETON WORD" AND "WRONGLY-SPELT WORD" TYPES OF TEST.

So far, both the W.Sp. and Skel. tests seem to be very satisfactory, and there is little to choose between them, although the W.Sp. has a slight

¹ *Loc. cit.*

² *Loc. cit.*

³ *Loc. cit.*

advantage. To make a choice, therefore, it is advisable to examine these two types of test more carefully, and find *how* the mistakes were made in each.

As one would expect from the nature of these tests, some of the "errors" consisted either of (a) *omissions* (which, since there was no time limit, must have been most often due to a pupil's not recognising what the word was which he was required to spell) or of (b) *misunderstandings*, i.e., writing a different word from that required (this also is the result of non-recognition). The following table gives the number of omissions and misunderstandings in the Skel. and W.Sp. tests.

TABLE V.
OMISSIONS AND MISUNDERSTANDINGS IN THE "SKELETON WORD" AND
"WRONGLY-SPELT WORD" TESTS.

	Skel.		W.Sp.	
	No.	Per cent.	No.	Per cent.
No. of omissions (out of 2,000 responses)	95	4.8	40	2
No. of misunderstandings (out of 2,000 responses)	31	1.6	7	0.4
TOTAL	126	6.3	47	2.4
No. of responses omitted which were correct in original dictation	29	1.5	10	0.5
No. of responses "misunderstood" which were correct in original dictation	9	.5	0	0
TOTAL	38	1.9	10	0.5

Thus, out of 2,000 responses in each case, although 126 errors in the Skel. and 47 in the W.Sp. test, were made through omissions and misunderstandings, most of these responses had been wrong in the original dictation too, and the actual *loss* of marks in this way was only 38 (out of 2,000), or 1.9 per cent, in the one case, and 10 (out of 2,000), or 0.5 per cent, in the other. Small as these figures are, however, they again favour the W.Sp. test.

The next table has reference to the W.Sp. tests alone, and shows the extent of the *disturbing influence* of the wrongly-spelt words which are given. (For example, when a pupil, who spelt "carriage" correctly in the original dictation, wrote "carridge" in the W.Sp. test, he was obviously influenced by the given spelling "caridge.")

TABLE VI.

EXTENT OF "DISTURBING INFLUENCE" IN W.SP. TESTS.

	No.	Per cent.
No. of responses (out of 2,000) wrong in the W.Sp. tests but correct in the original dictation.....	177	8.9
No. of responses, out of these 177, which show disturbing influence of wrong spellings given	92	4.6

(NOTE.—The phrase "out of these 177" is meant to draw attention to the fact that only those responses which were wrong in the W.Sp. tests but correct in the original dictation are being counted. Words wrong in both are not considered, even if, as sometimes happened, the wrong spelling in the W.Sp. test was different from the wrong spelling in the original dictation and showed the "disturbing influence" of the wrong spelling given.)

Table VI reveals "disturbing influence" as a fault in the W.Sp. test. 92 marks were lost because pupils, able to write down the word correctly when it was dictated, were misled by the wrong spellings confronting them. As I said earlier, I think that these pupils are just as good spellers as those who are able to resist the influence of the wrong spellings, and yet they are here placed at a disadvantage. It may happen, moreover (although I do not stress this), that these wrong spellings affect the child's spelling after the test is over. It seems even better that a mark should be lost by the omission of a response altogether than that it should be lost through the upsetting influence of a wrongly-spelt word.

X.—DECIDING BETWEEN "SKELETON WORD" AND "WRONGLY-SPELT WORD" TYPES OF TEST—POINTS FAVOURING FINAL SELECTION OF "SKELETON-WORD" TYPE AS THE BEST.

So far it has appeared that both Skel. and W.Sp. are good tests, but that both have certain disadvantages. The following additional considerations will help to swing the balance in favour of the Skel. test:

(a) If a few more letters were introduced here and there into the "skeleton words" used in the experiment, there would not be so many omissions and misunderstandings. The scores would be slightly higher, but the correlation with dictation would probably remain as good.

(b) Many of the omissions and misunderstandings were no doubt due to the pupil's being unfamiliar with the word in question. Ability in English, however, includes knowledge of meanings of words, and so, if a "skeleton word" spelling test were included in an English test, the few marks which might be lost otherwise than by inability in spelling would be lost by inability in another branch of English. The result would therefore not be impaired.

(c) "Disturbing influence," discussed above, is a serious, if not an extensive, fault in the W.Sp. test.

(d) The number of omissions and misunderstandings could be reduced by avoiding certain words which do not readily lend themselves to this type of test (e.g., the word "heir": if any more letters than the final "r" are given, too much information about the spelling of the word is given away).

(e) Apart from experimental evidence altogether, it is only reasonable to expect that a test of the "skeleton word" type would be the best, *provided that omissions and misunderstandings could be eliminated*. All that is required in a non-dictated test is something to enable the pupil to know easily and clearly what the word is that he is to write, without assisting him to spell it on the one hand, or hindering him from spelling it on the other. A careful consideration of the letters to be given and the lengths of spacing between the letters would surely make the "skeleton word" test as efficient as a spelling test can be.

XI.—SUMMARY OF RESULTS AND CONCLUSIONS.

(1) All three types of non-dictated spelling test are shown to be valid as measuring spelling ability.

(2) In the "Wrongly-Spelt Word" tests there were more, and in the "Multiple Choice" tests considerably fewer, errors than in the dictation, while the "Skeleton Word" tests showed about the same number as dictation.

(3) "Wrongly-Spelt Word" and "Skeleton Word" tests showed a greater number of responses than did "Multiple Choice" tests which were "similar" to their corresponding responses in dictation (i.e., both right or both wrong.)

(4) The correlations of the tests with the dictation criterion are: "Multiple Choice," .87; "Skeleton Word," .90; "Wrongly-Spelt Word," .91.

(5) From the above we see that "Wrongly-Spelt Word" and "Skeleton Word" tests are superior to "Multiple Choice" tests.

(6) It is observed that in 4·6 per cent of the cases pupils who had spelt a word correctly in the dictation spelt it, in the "Wrongly-Spelt Word" tests, wrongly and in such a way as showed the disturbing influence of the wrong spelling given.

(7) The only fault in the "Skeleton Word" tests is that responses were occasionally omitted or misunderstood. Remedies for this are indicated.

(8) All things considered, the "Skeleton Word" test is recommended as the best.

I wish to make grateful acknowledgment of a grant from the Thomson Research Fund, Department of Education, University of Edinburgh, for the purposes of this experiment, and to thank Mr. W. G. Emmett, of the same department, for his invaluable assistance and expert criticism during the preparation of this paper. I should also like to express my thanks to Mr. G. Mowat, Headmaster, Demonstration School, Moray House Training College, Edinburgh, and to Mr. F. Hamilton, Headmaster, Abbeyhill School, Edinburgh, for assisting me by placing certain classes in their schools at my disposal.

APPENDIX I.

TABLE SHOWING DISTRIBUTION OF AGES OF PUPILS.

<i>Age (in years and months).</i>	<i>Frequency.</i>	<i>Age (in years and months).</i>	<i>Frequency.</i>
13; 4 —13; 8	1	11; 3 —11; 7	6
12; 11—13; 3	2	10; 10—11; 2	6
12; 6 —12; 10	8	10; 5 —10; 9	10
12; 1 —12; 5	24	10; 0 —10; 4	15
11; 8 —12; 0	8		80

APPENDIX II.

COMPLETE LISTS OF WORDS USED IN THE EXPERIMENT.

<i>List A.</i>	<i>List B.</i>	<i>List C.</i>	<i>List D.</i>
Tuesday	geography	glue	British
telephone	carrot	linen	tease
fashion	enemy	paid	yawn
conquer	knowledge	cushion	quiet
convey	grammar	library	sandwich
piano	annoyed	echo	theatre
traveller	saucer	disguise	loose
swollen	Spaniard	spectator	quantity
coffee	ceiling	ancient	biscuit
envelope	quarrel	precious	bicycle
disease	electricity	succeed	therefore
choir	current	whose	criminal
heir	medicine	stomach	league
view	strength	volcano	carriage
deny	lightning	hymn	assistance
parliament	typewriter	interrupt	receive
author	accustomed	debt	tobacco
obedient	cupboard	believe	seize
breadth	unusual	yacht	foreign
muscle	Egypt	slaughter	convenient
excellent	extremely	horizon	calendar
niece	disappoint	precipice	peculiar
delicious	awkward	breathe	receipt
behaviour	handkerchief	sincerely	immense
paraffin	separate	embarrass	necessary

APPENDIX III.

TABLE SHOWING DISTRIBUTION OF SCORES IN THE ORIGINAL DICTATION.

<i>Scores (out of 100).</i>	<i>Frequency.</i>	<i>Scores (out of 100).</i>	<i>Frequency.</i>
90—99	3	50—59	8
80—89	10	40—49	9
70—79	19	30—39	7
60—69	16	20—29	8
			80

Résumé.UNE EXPÉRIENCE AVEC DES ÉPREUVES D'ORTHOGRAPHE
NON DICTÉES.

Le but de cette expérience était de calculer la validité de trois types d'épreuves d'orthographe non-dictées (appliquées par le sujet lui-même), c'est à dire : (1) "Le Choix Multiple" (l'orthographe correcte est à choisir entre plusieurs) ; (2) "Le Mot Incomplet" (quelques lettres seulement du mot sont données ; le mot est à compléter) ; (3) "La Faute d'Orthographe" (faute d'orthographe à corriger). Dans chaque cas l'on indiqua le contexte du mot. L'on adopta la dictée comme critère. Après avoir été dictés, les mêmes mots furent présentés aux mêmes élèves (par une méthode rotatoire) au moyen des épreuves non-dictées, et l'on compara les résultats avec ceux de la dictée.

L'on trouva que tous les trois types d'épreuves avaient une valeur, mais que celui du "Choix Multiple" en avait la moindre. Le deux autres étaient presque égaux, mais la valeur de la "Faute d'Orthographe" était diminuée par le fait que les fautes exerçaient quelquefois une influence mauvaise sur l'orthographe des élèves qui savaient épeler les mots dans des circonstances normales.

A tout prendre, le type d'épreuve du "Mot Incomplet" fut recommandé comme étant le meilleur.

ZUSAMMENFASSUNG.

EIN VERSUCH MIT NICHTDIKTIERTEN BUCHSTABIERTESTS.

Man versuchte, die Zuverlässigkeit von drei Arten von nichtdiktiertem ("selbstprüfendem") Buchstabiertest zu ermitteln, nämlich :

- (1) Test "Vielfache-Wahl" (die richtige Form soll aus mehreren herausgesucht werden).
- (2) Test "Skelett-Wort" (nur einige Buchstaben des Wortes werden gegeben : das Wort soll ganz geschrieben werden).
- (3) Test "Falschgeschriebenes-Wort" (das falschgeschriebene Wort soll richtig geschrieben werden). In jedem Fall wurde der Zusammenhang gegeben.

Das Diktat galt als Kriterium. Nachdem die Wörter diktirt wurden, wurden sie denselben Schülern durch Rotation vermittelt nichtdiktierter Tests gegeben, und man verglich die Resultate mit dem Diktat.

Alle drei Testtypen zeigten sich als gültig, von den dreien aber war der "Vielfache-Wahl"-Typus am wenigsten gültig. Die anderen zwei waren mehr oder minder gleichwertig, aber der Wert des Tests "Falschgeschriebenes-Wort" wurde durch die Tatsache gemindert, dass die falschgeschriebenen Wörter auch die Orthographie der Kinder manchmal störend wirkten, die sonst gut im Buchstabieren waren.

Alles in allem ist der Test "Skelett-Wort" als der beste zu empfehlen.

THE RELATIONS OF EDUCATIONAL ABILITIES.

By CYRIL BURT.

- I.—*Problem.*
- II.—*Data.*
- III.—*Methods of factor analysis.*
- IV.—*Comparison of results.*
- V.—*Stability of factors.*
- VI.—*Summary.*

I.—PROBLEM.

THE recent growth of educational and vocational guidance has done much to revive interest in 'mental types' and in 'special abilities'; but, in their views about such conceptions, the practical psychologist and the theoretical investigator seem to have drifted into opposite camps. Twenty years ago, the school psychologist was mainly concerned with classifying his individual cases according to a single 'general factor' of intelligence—with diagnosing mental defectives whose ability was subnormal in every direction, or with selecting for secondary schools and for higher professions children whose ability was supernormal in every direction. In those days, therefore, he was quite content with a theory that emphasized the ubiquity of all-round ability, and left no room for faculties or types. From that time onwards the orthodox view has apparently been that there is little evidence either for special abilities or for special disabilities. The few early efforts to demonstrate the presence of mental factors peculiar to broad groups of school subjects, or to broad classes of intellectual operations, were either condemned as belated revivals of the old faculty hypothesis, or at best dismissed as inconclusive because of the fewness of the cases and the crudity of the statistical procedure.

During recent years, however, the work of the new psychological clinics, which from their very nature have continually to deal with the rarer and more exceptional types of pupil, has been gradually forcing on our notice the existence of children with limited or specialized defects, often in a most extreme and unexpected form, side by side with the far commoner type, namely, those who are dull and backward in all directions. Nor is the vocational psychologist any longer content to grade a batch of applicants by tests of general intelligence alone: his reviews of industrial failures and successes have led him likewise to believe in the value of

special aptitudes or gifts. At the same time, within the ordinary school, the smaller classes and the more flexible time-table are permitting teachers to classify pupils differently for different subjects; and many are beginning to ask what subjects go most naturally together. Abroad, the notion of a multiplicity of 'common factors,' some more general and some less, which would largely justify a belief in such types and such classifications, has lately attracted an increasing measure of attention, particularly in America. It seems, therefore, high time that the theoretical psychologist should take up the whole question afresh.

A brief glance at the earlier results and the criticisms urged against them will perhaps enable us to understand more clearly the special difficulties that the investigator has to meet. With these in mind, we may be able to avoid the fallacies associated with the old faculty hypothesis without at the same time discarding the germ of truth that it seems to contain.

In a series of inquiries,¹ begun at Liverpool in 1909, I endeavoured to study the relations between pupils' attainments in the various subjects of the school curriculum and to discover by a statistical analysis what were the chief mental qualities affecting their progress, or lack of progress, in particular directions. With the generous aid of Miss M. E. Bickersteth, Mr. J. A. Davies, and Mr. R. C. Moore—all teachers with special training in the administration of psychological tests—numerous samples of children were tested, of both sexes, of different ages, and attending various types of school, and methods of analysis worked out. Shortly afterwards, when the post of educational psychologist was created by the London County Council, it proved possible to continue these surveys on a much larger scale, and to publish detailed tables in the Council's official *Reports* instead of leaving them buried in inaccessible theses.² In the largest inquiry of all, 120 children aged 10 to 12 were examined by means of standardized tests; correlations were calculated; and the correlation table subjected to a factor-analysis by methods which were admittedly but tentative, but appeared to lead to important new results. Teachers' mark lists were also analysed by the same devices to show that the tests (a novelty in those days) did not deal with mental processes foreign to the regular work of the classroom.

Our main conclusions were as follows. Educational attainments were "considered to depend upon factors of three different orders: (1) first, a

¹ BURT AND BICKERSTETH: *Some Results of Mental and Scholastic Tests: Report of Conference of Educational Associations* (Jan., 1916), pp. 30 et seq.; *L.C.C. Report on the Distribution and Relations of Educational Abilities* (1917), pp. 51-64. (For brevity I shall cite this latter publication as my '1917 Report'). The various statistical methods (several were tried out and not a few rejected) were described more fully with the necessary proofs in degree theses by the investigators named in the text.

² I am indebted to the Council for permission to publish the data quoted in the following article; and at the same time I should like to acknowledge the great service which the London Education Authority performed, at a time when educational research was widely regarded as a passing pedantry or fad, in encouraging systematic investigations by teachers and others, and in printing costly volumes of statistical tables and reports, which in those days no author or journal could have afforded to publish.

hypothetical *general* factor, entering into all the tests": (this was identified, not with 'general intelligence' or 'g', but rather with a wider or more complex factor which I called 'general educational ability'); "(2) secondly, a multiplicity of special or *group* factors, confined (at any rate so far as their positive saturation coefficients were concerned) to limited groups of subjects; (3) thirdly, specific, individual, or *unique* factors, peculiar to each particular test itself." This view was at the time described by one of our collaborators as a 'three factor theory', to distinguish it from the 'two factor theory'¹ which maintained that each test-performance could be analyzed into two components only, namely, the general factor (g) and its specific or unique component (s).

"The 'special' or 'group factors' are shown to be the following: (i) an arithmetical factor; (ii) a manual factor; (iii) a linguistic or verbal factor, which may perhaps be separated into two types, (a) an elementary or verbal (as in reading and spelling) and (b) a literary or informational (as in composition and essay subjects)." Corresponding to these, it was suggested, were certain convenient sub-classifications of children, justifying to some extent the practical teacher's belief in so-called 'types,' and grouping the more anomalous cases into instances of specific disability or specific ability in this or that limited direction.²

¹ Both labels are in my view apt to be a little misleading: they emphasize minor differences at the cost of the far more important points of agreement. As was argued at the time: "the so-called general factor is simply the particular group factor that is of most widespread range (i.e., highest variance) in the table under examination; the group factors are merely general factors some of whose test-saturations are indistinguishable from zero; and the specific factors may be simply group factors that happen to be limited in that table to one particular test." My own theorem (if it requires any label) should rather be called a 'four factor theorem': since among the 'unique' factors there are really factors of two kinds: first, the 'factor' specific to the test, and, secondly, the 'factor' representing chance errors, i.e., the test's 'unreliability or inconsistency' (see *Marks of Examiners*, table on p. 259, and the schematic factor-pattern in table on p. 264; also eq. xvii, p. 275).

² *Psychological Tests of Educable Capacity* (1924), p. 19; cf. *1917 Report*, p. 58. The existence of an arithmetical factor (curiously enough not found by Dr. Carey) was confirmed by subsequent work (e.g., D. J. Collar, *Brit. J. Psych.*, XI, 1920, pp. 135-158; A. Rogers, *Columbia University Contributions to Education*, 1923, No. 130). The existence of a verbal factor—owing perhaps to its apparent disproof by C. M. Davey (*Brit. J. Psych.*, XVII, 1926, pp. 27-48)—has only recently been accepted in this country: in America Kelley adduced considerable evidence in its support (*Crossroads in the Mind of Man*, 1928, pp. 104 *et seq.*), and more recently still Stephenson seems finally to have corroborated my original deductions (this *Journal*, XXII, 1931, pp. 255-67). The existence of a manual or practical factor has proved still more elusive; but strong arguments in its favour have been advanced by Alexander on the basis of very extensive tests: it may be added that he also finds evidence for the other factors—general, verbal, and numerical: (*Intelligence: Concrete and Abstract*, *Brit. J. Mon. Sup.*, 1935, pp. 95-99, 86). The occurrence of 'types' of children specifically limited in the 'verbal' and the 'numerical' factors respectively has been carefully investigated by two of my own research students—Miss Gertrude Hume, formerly Lecturer at Graystone Place Training College, and Miss Elizabeth Wheeler, L.C.C. Inspector of Schools. By means of questionnaires addressed to London children, followed by an intensive study of the individual cases thus revealed, Miss Hume verified the existence of children specifically backward in reading and spelling, and Miss Wheeler that of children specifically backward in arithmetic (see their unpublished university theses); but they find that genuine and unmis-takeable cases of the two extreme 'types' are decidedly rarer than has been supposed. The 'tendencies,' however, are common enough, as is easily verified by correlating persons.

Of this conclusion, the first half was immediately accepted, and the other half repeatedly challenged. The presence of a 'general factor', together with the practical consequences that it entails (the feasibility of examining children for all-round abilities and of classifying them accordingly, whether in separate classes or in separate schools), and the inevitable occurrence of specific factors peculiar to each particular test—these are points that have hardly been questioned. On the other hand, the suggestion of additional 'group factors' was received at first with considerable criticism and doubt.¹ Professor Ward declared them to be a "dubious resurrection of the long discarded doctrine of faculties and types." Professor Spearman was far more sympathetic. Yet, only ten years ago, after a review of all the available evidence, he found himself obliged to admit that the final conclusion was "still negative." "Cases of specific correlations or group factors," he writes, "have been astonishingly rare. Of special abilities sufficiently broad to admit of measurement, there have been but the scantiest indications. Thus the modern version of the doctrine of faculties is none the happier for discarding the old name while retaining the old fallacy."²

The difficulty has always been that the amount of variance attributable to group factors is admittedly small. Hence those who believe in such factors can only bring a convincing case when they have first accumulated data from numbers large enough to reduce the probable errors to very slender proportions. In the early days of educational research, individual tests alone were used; and batches of 12 to 30 children only were tested. Under such conditions minor factors, each responsible for less than 10 or 15 per cent of the total variance, were inevitably eclipsed by the main general factor which is usually responsible for at least 40 or 50 per cent. When written or 'group' testing was introduced, large numbers of children could be tested in the same experiment; partial correlation could be legitimately applied; and the small residuals began to claim statistical significance.³ As soon as this step was

¹ In my previous research on "Experimental Tests of General Intelligence" (*Brit. J. Psych.*, III, 1909, p. 164) I had found some slender evidence for group factors of a more purely psychological kind—viz., in memory tests, in sensory tests, and in motor tests respectively; and I argued more fully for the recognition of such factors in a paper read to the Manchester Child Study Society in 1909 (reprinted in *Child Study*, Vol. IV, Nos. 2 and 3, see esp. pp. 94 and 95). Later Dr. Carey, then a colleague and a teacher in the Council's service, also adduced evidence from teachers' marks pointing in a similar direction: her results, together with a summary of my own conclusions, more accessible to the psychological reader, will be found in *Brit. J. Psych.*, VIII, 1916, pp. 176-180. Professor Godfrey Thomson does not reject the notion of a 'general factor' (defined as a merely descriptive or practical conception, virtually signifying 'the average ability which a man shows') but he is 'not a believer in a faculty called general ability' with which the general factor is commonly identified (cf. *Tests of Educable Capacity*, pp. 234, 236).

² SPEARMAN: *The Abilities of Man* (1927), p. 241.

³ The value and reliability of group testing in attempting to measure the intellectual abilities and attainments of school children were demonstrated by Moore, Davies and myself in a series of researches made possible by the generous co-operation of the Liverpool Education Authority and summarized in the first volume of this *Journal* (then issued as the *Journal of Experimental Pedagogy*: see esp. I, 1911, pp. 93-112, 233-388). These were, we believe, the first investigations in this country in which the number of children tested ran into well over a hundred. I readily agree that the results of group testing always require checking by individual testing, if possible with introspections. I note, too, that American opinion is now inclined to regard group testing as suited only for the first broad statistical surveys (cf. Thurstone, *Primary Mental Abilities*, 1938, p. vi).

taken, it became evident that the single factor theory must give way to some kind of multiple factor theory. But if as many as four or five factors are to be conclusively established, the number tested must amount to nearer a thousand than a hundred.

II.—DATA.

In the course of twenty years' work for the London County Council I have been able to accumulate far more extensive measurements by means of standardized tests. Here I can only take up one of the many problems which they raise. To include differences attributable to difference in sex or age would obscure the main issue and at the same time take us too far afield. Accordingly, I shall confine myself primarily to data from boys belonging to a single age-group, those aged 10 last birthday: of these, 613 have now been tested along comparable lines.

For most of the school subjects, the tests used were similar to those described in one of my later L.C.C. *Reports*¹; for history, geography, science or nature study, handwriting and drawing, additional tests were employed, varying to some extent with the differences in the curricula at the different schools.

The observed correlations for the thirteen school subjects are set out in Table I. If the reader compares them with tables already published² he will see that the present survey shows coefficients of much the same magnitude and arrangement as those previously obtained both from the results of my own tests and from the school marks assigned by teachers.

III.—METHODS OF FACTOR-ANALYSIS.

For analysing such a table of correlations various methods have recently been put forward; and it is popularly supposed that the current disagreement over methods entails an equal disagreement over results. My first object, therefore, will be to show that, with the present data at least, each method leads to much the same conclusions, and that each is broadly in keeping with the analysis originally proposed. Since the coefficients in Table I, like those in the earlier table, are all without exception positive, we are tempted to explain the correlations by the presence of common factors. But such an explanation at once brings with it a further question: 'Is one common factor enough? Or' (to quote the words of the *Report*) 'have we to recognise a multiplicity of common factors, each limited to certain groups of subjects?'³ To decide

¹ *Mental and Scholastic Tests*, 1921, Memorandum III, pp. 339-369.

² Table I in *Some Results of Mental and Scholastic Tests*; Tables XVIII and XXII in the 1917 *Report*. It should be noted that in the present table the bracketed figures given for the self-correlations in the leading diagonal are not 'reliability coefficients,' but the total variances attributable to significant factors.

³ 1917 *Report*, p. 53.

TABLE I.
CORRELATIONS BETWEEN SCHOLASTIC TESTS.

Test.	Composition.	Reading (Compre- hension.)	Reading (Speed.)	Dicta- tion.	History.	Geo- graphy.	Science.	Totals of Sub- matrices.	Arith- metic (Prob.).	Arith- metic (Rules).	Totals of Sub- matrices.	Handi- work.	Drawing.	Writing (Qual.).	Writing (Speed).	Totals of Sub- matrices.
Composition	—	.723	.429	.598	.677	.612	.461		.578	.283		.277	.263	.231	.205	6.821
Reading723	—	.346	.516	.506	.484	.426		.378	.134		.268	.198	.272	.197	
Comprehension429	.346	—	.303	.337	.329	.242		.313	.151		.253	.254	.229	.142	
Reading (Speed)598	.516	.303	—	.414	.410	.430		.289	.066	4.011	.274	.162	.297	.173	
Dictation677	.506	.337	.414	—	.510	.262		.535	.308		.326	.365	.261	.218	
History612	.484	.329	.410	.510	—	.281		.482	.257		.371	.396	.319	.237	
Geography461	.426	.242	.430	.262	.281	—		.226	.011		.281	.037	.290	.025	
Science																
Arithmetic (Prob.)									—	.561		.468	.548	.279	.108	2.185
Arithmetic (Rules)561	—		.262	.389	.097	.034	
Handwork																
Drawing																
Writing (Qual.)																
Writing (Speed)																
Totals of Submatrices																
Arithmetic (Prob.)																
Arithmetic (Rules)																
Handwork																
Drawing																
Writing (Qual.)																
Writing (Speed)																
Totals of Columns with diagonal submatrices omitted	1.837	1.447	1.322	1.261	2.033	2.062	.870	—	4.204	1.962	—	2.780	2.612	2.275	1.339	—

the issue it was suggested that we might take, as a criterion, the difference between the coefficients actually observed and the theoretical coefficients that could be deduced on the assumption of one general factor only. If this difference was too large to be explained by the chances of sampling, then, it was argued, we must seek additional factors. Actually, when the probable error was sufficiently reduced by increasing the numbers tested, there seemed no doubt whatever that a single general factor was inadequate to account for the total correlations observed.

The next problem was to determine the location of these additional factors. Various tentative methods were proposed, all based upon the following principle. A single factor, conceived as entering into a series of tests and so determining their inter-correlations, tends necessarily to produce what is termed a hierarchy, i.e., an arrangement of coefficients in which each row is proportional to every other. When a single factor is no longer enough to account for the figures actually observed, I suggested that we might then regard the observed table not as a poor approximation to a single hierarchy, but as *the sum of a number of hierarchies superposed*, each hierarchy being the result of its own particular factor.¹

Such superposed hierarchies, however, might either cover *all* the tests in the table or be limited to certain *groups* or clusters of tests. In the former case any additional factor could be still called a 'general' factor (so that there might be more than one); in the latter case it was described as a 'group' factor. These 'general' and 'group' hierarchies were illustrated in Tables XVIII and XX of my 1917 *Report*. The two possibilities, it was added, necessitated 'slight but obvious complications in the summation formula' employed to determine the factor saturations; and, as further studies have shown, the alternative results to which they lead seem at first sight to be somewhat divergent. Before going further, therefore, it appears desirable to consider a little more fully the differences between the two methods and the assumptions they imply.

(a) *Group Factor Methods.*²

The ordinary notion of a group factor rests on a double assumption: it assumes (i) that such a factor is by definition sharply limited to a certain

¹ I have attempted a more rigorous formulation of this fundamental principle and its corollaries in *Psychometrika*, III, pp. 151-168 ('The Unit Hierarchy and its Properties').

² The methods described in the following paragraphs were summarized in the original draft of my Memorandum on factor-analysis, prepared in 1935 at the request of the Secretary of the International Institute Examinations Inquiry and subsequently published in the Institute's volume on *Marks of Examiners* (Memorandum I, pp. 254-314; hereafter cited as '1935 Memorandum'). The account of the group-factor method (pp. 306 *et seq.*) was considerably abridged, as the data did not seem adequate for determining such additional factors. The procedure had already been used by research students and others, and described, with worked examples, in several of the theses already cited. The available formulae have been brought together in *Journ. Psych.*, IV, pp. 347-356. The 'general factor methods' (Memorandum, pp. 284 and 287, *et seq.*) have been more fully discussed in a recent number of this *Journal* (VII, pp. 172-195).

group of variables ; and (ii) that its saturation coefficients for those variables are all positive. The effect of such a group factor is thus to add a positive specific correlation to any pre-existing inter-correlations between the variables within the group, but zero correlation elsewhere. The natural test for these assumptions will therefore be (after first eliminating the general factor, if any) to re-arrange the table so that the whole matrix can be partitioned into sub-matrices, in correspondence with the divisions between the group factors ; the dividing lines should then mark out (a) square blocks of positive specific correlations, lying along the leading diagonal, all the correlations in one block being affected by one and the same group factor, and (b) rectangular matrices elsewhere containing zero specific correlations, or at least correlations fluctuating on either side of zero within the limits imposed by the probable error.

This principle was adopted in my original research ; and illustrative examples of the rearrangement were given in Tables XX and XXIII of the *Report*.¹ The resultant grouping of the subjects was exhibited by a clock-diagram (*loc. cit.*, Figure 8) showing the position of each test on the circumference of the section of a sphere, and demonstrating the bunching of related subjects into four distinguishable clusters. A separate factor was located at the centre of each cluster. But, as was remarked at the time, the linking of the traits on the outer edges of adjacent clusters rendered the precise delimitation of the different factors somewhat doubtful : thus, among the 'verbal' subjects; the more mechanical (reading and dictation) seemed to form one cluster, and the informational (composition, history, geography, and science) a second ; yet one of the reading tests was not so closely related to the other reading test as it was to composition : similarly, science fell almost midway between the informational group of subjects and the arithmetical group, and writing almost midway between the verbal group and the manual. This phenomenon (which reappeared in many other tables) I designated 'cyclic overlap' ; and it was noted that, if the 'group factors' themselves overlapped (as this feature seemed to imply), then the divisions between the various groups of tests could not be very sharply drawn.

These various considerations imply a number of requirements upon which any satisfactory analysis into submatrices and of submatrices must be based : (i) there shall be at most only one general factor ; (ii) the group-factors, like the general factor, must represent positive abilities, i.e., must always favour and never hinder actual test performances, and can therefore possess positive but no negative saturation coefficients ; (iii) these group-

¹ The same method had been described and used in a study of 'General and Specific Factors underlying the Primary Emotions' : (see tables and diagram published in *Brit. Ass. Ann. Report*, 1915, p. 695). In my later diagrams I substituted the principle described and illustrated by Maxwell Garnett in his re-examination of my results (*Brit. J. Psych.*, IX, pp. 150, 153, figs. 1 and 2). Somewhat similar clock-diagrams, showing clusters produced by 'verbal', 'practical', and other group factors, have since been published by Alexander (*Intelligence, Concrete and Abstract*, pp. 143 *et seq.*).

factors are by definition limited to certain tests or traits only, and must therefore possess zero saturations for all other tests or traits; (iv) the groups which the factors cover may at times overlap, i.e., two or more group-factors may occasionally enter into one and the same test; (v) there shall be, within the whole table of intercorrelations, at least as many submatrices as there are group-factors in which the intercorrelations themselves remain unaffected by the group-factors.

In dealing with any particular table, the first step will nearly always be to calculate the saturation coefficients for the first or general factor, and then by their means, eliminate its influence. Now if we adopt the usual procedure, and attempt to analyse the correlation matrix as a whole, the outcome will inevitably be a set of secondary factors running through all the traits, instead of through certain groups, and having negative as well as positive saturations, more or less equally balanced. To circumvent this difficulty, I proposed a 'submatrix method', in which we work neither with individual coefficients nor with the matrix as a whole, but with selected submatrices. Various formulae can be devised, yielding more or less satisfactory results. For the most part they depend upon the fact that practically all the well-known equations, usually expressed in terms of individual correlations, are equally applicable to the submatrices, and to the totals of the submatrices and of their constituent columns.

The observed correlations having been rearranged in the manner described above, with the coefficients affected by the f group-factors lying along the leading diagonal, the total matrix R is then, as we have seen, partitioned symmetrically into $f \times f$ submatrices. Adopting the conventional notation¹ for the minors of a partitioned matrix, we may express this algebraically as follows:

$$R = \begin{bmatrix} R_{11} & R'_{21} & R'_{31} & \dots & R'_{f1} \\ R_{21} & R_{22} & R'_{32} & \dots & R'_{f2} \\ R_{31} & R_{32} & R_{33} & \dots & R'_{f3} \\ \dots & \dots & \dots & \dots & \dots \\ R_{f1} & R_{f2} & R_{f3} & \dots & R_{ff} \end{bmatrix} \dots\dots\dots (i)$$

The submatrices $R_{11}, R_{22}, \dots, R_{ff}$, which break the hierarchy, can now be treated like the 'reliability coefficients' $r_{11}, r_{22}, \dots, r_{nn}$ in an ordinary correlation table, and either omitted or replaced by estimated totals which fit into the hierarchy. If we could assume that the surviving part of the table is *exactly* hierarchical, then we should be led, as I have shown elsewhere,² to equations for the saturation coefficients which in form are precisely analogous to those deduced by Spearman for the ordinary hierarchical table (*loc. cit.*, equations v and vi): these formulae were, as a matter of fact, adopted in some of our earlier researches. But when the hierarchical arrangement is approximate only, they would seem to be open to the same objections as their simpler analogues. The defects, however, can be largely overcome by using slightly more elaborate formulae.

¹ ATKIN: *Theory of Canonical Matrices*, p. 6. It is perhaps a little unfortunate that the same symbol, R_{ji} , is also used to denote the cofactor of r_{ji} .

² "Factor Analysis by Submatrices."—*Journ. Psychol.*, VI, pp. 339-375.

When, as here, three group-factors alone are discernible, the working formula reduces to the following expression¹:

$$r_{1g} = \frac{\sum_j R_{1j}}{\left\{ \sqrt{\frac{\sum R_{2j}}{\sum R_{31}}} + \sqrt{\frac{\sum R_{3j}}{\sum R_{21}}} \right\} \sqrt{\sum R_{1j}}}$$

where r_{1g} denotes the saturation coefficient for the first test, $\sum_j R_{1j}$ denotes the sum of the correlations remaining in the first column after R_{11} has been deleted, and $\sum R_{31}$ denotes the sum of all the correlations in the submatrix R_{31} (with analogous meanings for the other symbols and analogous equations for the other tests). The calculation is extremely simple: the only figures needed (all sums) are shown at the foot and sides of the submatrices in Table I. Thus, for Composition we have

$$r_{1g} = \frac{1.837}{\left\{ \sqrt{\frac{4.011}{6.821}} + \sqrt{\frac{6.821}{4.011}} \right\} \sqrt{2.185}} = .600$$

With correlation tables constructed artificially, this formula (unlike most of those in common use) will lead back to the saturation coefficients from which the artificial table was built up.

On analysing Table I by the method just described², we obtain for the first or 'general' factor the saturation coefficients shown in the first line of Table IIA. The relative magnitudes of these coefficients are much the same as those obtained in my 1917 research; the only important difference is that 'problem arithmetic' now shows a slightly higher correlation with the general factor than 'composition'.³ These saturation coefficients can next be used to eliminate this 'general' factor. We are then left with a set of residual correlations, in which most of the positive coefficients fall into three square submatrices or clusters, like those in my original *Report*. With the exception of the test for science, the residual correlations within the first group of tests (composition, etc.) show a fairly good approach to a hierarchical order: the correlations within the last group (writing, etc.) do not fit quite so well. Nevertheless, within each group the agreement is sufficiently close to justify us in treating these particular submatrices as if they were separate tables and extracting one factor from each: we then find that practically all the further residuals are devoid of statistical significance. The additional saturation coefficients thus obtained are shown in the remaining lines of Table IIA. We have, therefore, now reached a series of factors which fulfil the prescribed requirement, namely, that all should have either positive or zero saturation coefficients only.⁴

¹ *Loc. cit.*, p. 347, equation ii.

² I am especially indebted to Miss E. Carscallen and Miss G. Bruce for undertaking the greater part of the arithmetical work involved in these calculations.

³ A slightly different formula was used in Davies' thesis, where the 1917 table is analyzed by the group-factor method in fuller detail. The relative order, however, is still the same; i.e., in that research composition consistently gave higher saturations than problem arithmetic.

⁴ The factorial matrix thus obtained comes very near to what Thurstone describes as a 'simple structure' (*Vectors of the Mind*, pp. 150, 156, 166). The chief difference apparently is that Thurstone's requirements exclude the possibility that any single primary factor should be common to all the tests.

The psychological nature of the several factors can be readily inferred by noting the particular tests into which each enters. The first may be described as a factor of general educational capacity (using the word capacity in a somewhat loose sense), and the others as a verbal, a manual, and an arithmetical factor respectively. In short, the conclusions to be drawn from this new and more extensive set of data essentially confirm those of my original analysis twenty years ago. There is one slight divergence. In the earlier analysis, as we have seen, it was suggested that the verbal factor might be further sub-divided: there seemed to be first a factor concerned with the more mechanical elements of the verbal or linguistic subjects—particularly reading and dictation, and secondly a more complex factor connected with the higher processes of comprehension and expression through words, as in composition and the informational subjects. Here there seems no clear evidence for this further distinction, though on other grounds I still believe it to be valid.

It is important to be clear as to what precisely is meant by these 'special' or 'group' factors. Too often the teacher identifies the so-called 'verbal factor' with verbal performance in the concrete; and observes that the more intelligent child is also better in verbal work as he understands it. That, of course, is perfectly true: words express more general and more abstract ideas and relations; hence verbal work requires intelligence as well as special verbal capacity. But the group factors here contemplated endeavour to explain merely what can *not* be already accounted for by intelligence or general capacity, and that alone. Perhaps their nature will be clearer if we think of comparing and correlating, not tests, but persons. In such a case we either use tests of equal difficulty, or else first reduce the tests to equal difficulty by eliminating the general factor for persons. We then find that some individuals are better at the verbal tests as such than at the non-verbal as such, while others—often equally intelligent—are better at the verbal, *regardless of the amount of intelligence that either requires*. Thus, as was maintained in my *Report*,¹ these 'group-factors' are the factors that enable us to explain what are popularly called 'mental types.'

The theoretical table that can be constructed from the four sets of saturation coefficients given in Table IIA fits the original or observed correlations pretty closely. There are, however, a few instructive instances in which this artificial reconstruction deviates a little from the original. First of all, as in my earlier table, there are indications that the two writing tests have some small 'specific correlation', for the most part positive, with the verbal tests—composition, dictation, and reading: in one instance the correlation is negative, namely, between quality of writing and composition.² Secondly, drawing shows some correlation with the two arithmetic tests, and handwork with the tests of science and nature study: but this was only to be expected; 'geometrical drawing' was included, not only in some of the tests of drawing themselves, but also in the lessons on both drawing and geometry given at

¹ Pp. 64-5. The child's approximation to a given type is measured by his 'factor-measurement' (or, if we begin by correlating persons, by his saturation coefficient) for the corresponding 'special factor.'

² This was also one of the few specific correlations that proved to be both negative and significant in my earlier research: in explanation, I then suggested that "children who write fuller and better papers in examinations on written composition tend to scribble in a more hurried but less elegant hand than those whose answers are more meagre and whose attention is directed rather towards writing with a neat and careful hand" (1917 *Report*, p. 58).

certain schools; in the science lessons a good deal of construction and manipulation of apparatus was also introduced.

Such cross-correlations, however, are left out of account by the simple factor-pattern that we have just described. They imply an overlapping of the groups. So far as the positive residuals are concerned, these further peculiarities could, no doubt, be brought into the pattern by a slight complication in the procedure. For example, we could obviously make a fresh calculation after omitting all the overlapping tests, or again (what proves in the present instance to yield virtually the same result) we could choose appropriate tests as reference values.¹ The more important of the further figures obtainable in this way I have inserted (in brackets); but, owing to the fewness of the tests, such computations can claim but little value. Moreover, once these modifications are admitted into the simple non-overlapping scheme, it is not easy to see where to stop. With a large enough sample all our residuals could be rendered significant at almost any stage; and that in turn would mean that our 'group' factors would cease to be strictly 'group', since each would then be extended to cover every test.

The negative residuals may seem at first more problematical. They have, as Alexander puts it, "always given rise to considerable heartburning among psychologists." "It is difficult," he explains, "to interpret a negative correlation between a test and an ability."² Similarly, Thurstone argues that, "where psychological tests are involved, a negative factor loading implies an ability whose possession is a detriment to the test performance"; and that, together with the associated notion of a test that might be "facilitated by the lack of some ability," appears to him to be a contradiction in terms. Hence he too propounds the hypothesis that, where 'primary' factors are concerned, "the non-vanishing entries in the factorial matrix should all be positive."³

Now I venture to suggest that this postulate—that intelligible factors or components can have nothing but positive loadings or saturation coefficients—comes dangerously near an illegitimate identification of factors with faculties. As I have already insisted, a factor is, by definition and by derivation, a weighted average. It must represent, in the present condition of mental testing, hardly ever a single, simple ability, but nearly always a complex resultant of different ingredients—of acquired attainment as well as innate aptitude, of conative drive as well as sheer cognitive capacity.⁴ Whether or

¹ The appropriate formula is either equation xxvii (with the summation method) or equation xxviii (with the least squares method) as given in the 1935 *Memorandum*, pp. 286-7, or equation (viii) of *Journ. Psych.* VI, p. 355.

² *Loc. cit.*, p. 99.

³ *Vectors of the Mind*, pp. 161, 165-6. See also his chapter on 'The Positive Manifold' and his section on 'The Elimination of Negative Factor Loadings'; and his newer volume on *Primary Mental Abilities*, p. 71.

⁴ A careful reading of Thurstone and Alexander's arguments will, I think, reveal a tacit but erroneous assumption similar to that so often made by intellectualistic psychologists in the past (see more particularly *Vectors of the Mind*, p. 51, where 'ability' is defined but 'factor' is not, and *Intelligence: Concrete and Abstract*, pp. 3, 99). The assumption is that the empirical correlations between tests intended to measure 'abilities' can be actually analyzed into 'abilities' without remainder. Consequently on this assumption no room is left for any other mental operation—e.g., interference, inhibition, compensation, displacement of energy from one type of performance to another, and the like. In point of fact, however, as every teacher knows, good performance in (say) manual subjects is very frequently an indirect and compensatory result of lack of ability in other subjects, particularly the more formal (reading, writing, arithmetic, composition, etc.); still more often, the possession of high ability in one direction (e.g., English composition) may be detrimental to normal exertion in other directions (e.g., quality of handwriting or arithmetic).

not we believe that the 'general factor of intelligence' is mainly inborn, the more special 'abilities' must certainly depend quite as much upon interest and practice as upon any inherent gifts. Each pupil possesses only limited amounts of time and of energy. If he spends all his spare hours in reading, they will of necessity be unavailable for practising other subjects, such as handwork or arithmetic. If he devotes his attention chiefly to the more concrete subjects like handwork or drawing, then abstract subjects like arithmetic and reading are bound to suffer.

Thus a specific correlation which is found to hold good between (let us say) tests of arithmetic, after the influence of the general factor has been ruled out, may easily be due, not merely to the fact that certain children have a special talent for that subject, which helps them to do well in all arithmetical tests, but also to the fact that others heartily dislike it, or have not practised it so much, and consequently fail to exert or to reveal their full abilities in the tests. And generally a positive 'saturation' shown by a factor for one particular subject may quite intelligibly entail a negative 'saturation' for other subjects, which compete with it for the child's time and attention.

If the foregoing arguments be accepted, we must allow the secondary factors freedom not only to affect every test instead of only a small group, but also to show negative loadings as well as positive. But in that case the foregoing method of analysis ceases to be valid except as a first approximation: we must no longer look for one general factor and several group factors; we must look for a series of general factors. The apparent group-character of a factor will now be due, not to the fact that its saturation coefficients for all tests outside its group are strictly zero, but rather to the fact that, as contrasted with the large positive coefficients for tests within its group, the remaining coefficients, whether positive or negative, are relatively small.

(b) *General Factor Methods.*

The essential characteristic of a 'general' factor is that it constitutes an average (weighted or unweighted) of *all* the test-measurements in the table to which it relates, not merely of some. It follows that the simplest way of calculating its saturation coefficients for any one test is to average or sum the entire column of correlations furnished by that test, with or without an appropriate weighting. This principle leads to the two methods which I have already fully described in a previous number of this *Journal*¹—namely, what I called the simple summation method and the least squares method, the first forming merely a first (or unweighted) approximation to the latter.

When applied to a matrix of variances and covariances (including tables derived from measurements in standard measure, i.e., what would usually be

¹ VII, pp. 177-8. Formal proofs are given in my earlier *Memorandum* (pp. 284 and 291 respectively).

described as a matrix of correlations with unity in the leading diagonal), the results of the 'least squares method' yield a perfect fit both to the covariances themselves and to the original measurements; the 'simple summation method' under these conditions gives factor loadings that are in effect the first approximation to those given by the 'least squares.' Here, however, it is doubtful whether we should be justified in treating the variances of the different tests as equal. Certainly, in analysing an ordinary correlation table in terms of common or general factors only, it would at first sight appear illogical to make the saturations for the general factor cover factors that are specific to each test.¹ For our present purpose, therefore, the best plan is to assume n general factors for n tests, to be extracted in order of the contribution to the total variance, and to insert (so far as practicable calculation allows) the communalities arising from all those factors. Since these conditions will not by themselves lead to unique determinations for the communalities, we may seek more particularly saturations that (i) are deviations about their own mean as zero (except in the case of the first factor) and (ii) are statistically independent, i.e., uncorrelated one with another. In practice it is possible, as a rule, to realise these requirements only imperfectly, by means of successive approximation. We may attempt to approach the ideal results either by simple summation, or by the weighted summation method (table-by-column multiplication) associated with analyses by least squares. When the two conditions are realized precisely, the final outcome of both methods is the same.

Whether reached by simple or by weighted summation, the relation of the new general factors to the old group-factors will be fairly obvious. An average amount of each group ability will be transferred to the first or 'general' factor. The second and later factors will only express deviations from that average. Thus the child who is specifically of a verbal type (say) will be indicated by a plus measurement for the verbal factor, while the child who is of a non-verbal type will now be indicated, not by a zero but by a minus measurement. Each of the secondary factors, therefore, will now be bipolar, i.e., will have negative as well as positive saturation coefficients. We thus fulfil the first of our supplementary requirements: we can obtain and account for the negative residuals.

The second requirement was to make, not only the first hierarchy, but all the hierarchies cover the entire table. It has been assumed (e.g., by

¹ The treatment of group-factors suggested above implies that a specific factor may be regarded as a group factor peculiar to a group of one. If, as we shall see in a moment, group factors are converted into difference factors when a general factor method is employed, there seems no theoretical objection to according the same treatment to the specific. Here, however, the specific factors peculiar to each test (which of course also include their peculiar errors) lie outside our immediate field of interest.

I should add that, even when dealing with an ordinary correlation table, I still like to regard it as a covariance table. This means that I must treat the variances of the several tests as different and the communalities as providing an estimate of the variance. If the teacher were asked whether a random group of children vary more in Composition and Problem Arithmetic than they do in Speed of Reading or Writing, he would unhesitatingly answer yes. The need for assumptions (i) and (ii) is evident if we consider the analysis of correlations between persons: the factor saturations for persons are then analogous to the factor-measurements for those persons as obtained by correlating tests: these latter are always assumed to be uncorrelated and to be expressed as deviations about their own mean as zero.

other followers of Professor Spearman,¹ though not, I think, by Professor Spearman himself) that the general factor method cannot be applied afresh to the entire table of residual or specific correlations. Evidently, if we keep to simple unweighted summation, the positive and negative residuals, being in effect deviations about an average, must now add up to zero ; so that the summation formula cannot be applied without some complication in the procedure. Consequently it has been inferred that, when a pattern of positive and negative correlations appears, the group factor method is alone available. From this it would follow that, for any given table, there can be but *one* general factor.

But when we come to reconstruct our hierarchies it appears at once that something is wrong with this assumption. If, for example, we take a single row of saturation coefficients, and make some negative and others positive, we shall at once obtain a correlation table that forms a perfect hierarchy, and obeys the tetrad difference and every other test for a matrix of rank one. Moreover, if instead of the summation method we apply the least squares method to an empirical matrix of bipolar form, and proceed by successive approximation, then no matter what figures we take for our trial weights we are ultimately led to a positive and a negative loading of the antagonistic tests. Consequently, if we regard the summation method as in some sense an approximation to that of least squares (i.e., as an essentially analogous procedure except that the absolute values of the weights are unity), then we can legitimately adopt the same principle and prefix appropriate positive and negative signs to the unit weights.

Once again, therefore, we may begin by rearranging the table of residuals according to the method described for demonstrating 'cyclic overlap.' The result will exhibit a broad resemblance to a bipolar hierarchy, with negative coefficients for the most part massed together in two corner submatrices that are diagonally opposite. We have now merely to reverse the signs of the measurements for those tests which evidently have negative saturation coefficients ; and that, of course, will entail corresponding reversals in the signs of the related correlations.

We can then proceed as before ; and the result will be a further factor extending over the whole table—i.e., another 'general' factor.

The saturation coefficients obtained by simple summation are shown in the Table (IIc). The coefficients for the first factor now correspond rather more closely with those given in my earlier table in the 1917 *Report*. With the present table, too, as with that, the elimination of the first factor now leaves a number of significant negative residuals, as well as positive. In the earlier research, it was found, 'out of 78 coefficients only 25 are significant, but 9 of the significant coefficients are negative' ; and, as was then pointed out, these for the most part appeared just at the points where teachers commonly believe in 'antipathetic relations.'² In the present research the probable error has been reduced so much that

¹ Cf., for example, Stephenson's discussion of what he has termed 'system 5' (*Brit. J. Psych.*, XXVI, 1936, pp. 348-9, 351).

² *Loc. cit.*, pp. 57, 58.

as many as 51 of the residuals appear to be significant, about half being negative. Accordingly, we may now proceed to partial out, not only the first factor, but also a second, and even (it would seem) a third and a fourth.

In passing, it may be noted that the summation formula used for this purpose (and cited in my 1917 *Report*) is identical with that subsequently proposed by Thurstone. We may, therefore, regard the figures given in Table IIc as virtually¹ identical with those that would result from his 'centroid method,' before any supplementary rotation has been attempted.

On seeking to reconstruct the original correlations from these four general factors, we reach a somewhat better fit than that supplied by one general factor and three group. We may, however, inquire whether this is really the best fit that can be secured with this number of factors. The hypothetical matrix of rank f (where f denotes the number of factors) which forms the best approximation to a given correlation matrix is, as I have endeavoured to show in my 1935 Memorandum, the matrix reconstructed from the first f rows of the saturation coefficients obtained by the least squares method, i.e., from the first f factors extracted in order of their maximum contributions to the total variance. For the sake of comparison, therefore, an analysis by the least squares method has also been undertaken: the saturation coefficients obtained are appended in Table IIId.²

¹ The chief differences seem to be the following. In theory, Thurstone proposes to insert communalities that will give the lowest possible rank to the initial matrix (and therefore the fewest possible factors) within the margin allowed by the sampling error. This seems rather like assuming that the most probable difference between two groups (two mental types or the two sexes, for example) is the smallest difference compatible with the sampling errors: and this in turn would mean subtracting from each group-average, not the most probable error (which is zero), but the largest that convention permits. I, on the other hand, have assumed that an ideal analysis should completely account for the whole matrix as it stands. Actually I calculate detailed saturation coefficients only for those few factors that are statistically significant; nevertheless, as it seems to me, the influence of the remainder should be added to the communality, otherwise we do not obtain the most probable values. In practice Thurstone assumes that a fair estimate for the communality is given by the highest correlation in each column—an assumption which is quite unsafe with small tables (cf. *J. Psych.*, VI, p. 367); I insert a provisional figure obtained by mental smoothing, and then readjust by successive approximation. Finally, Thurstone's method appears, as a rule, to lead to factors whose saturations have an appreciable correlation: whereas, on my principles, the correlation should ideally be zero.

² Cf. *Memorandum, loc. cit. sup.*, pp. 286, 288 *et seq.* The figures are not those that would be obtained by Hotelling's method, since he inserts unity in the principal diagonal, whereas I insert the full communalities only (a procedure which gives a better fit to the observed correlations when only a few factors are to be extracted). In practice the quickest procedure is to start with the communalities and the saturations reached by 'simple summation' and then continue with 'weighted summation,' i.e., table-by-column multiplication (see this *Journal*, VII, p. 188, Table V): revised communalities are then inserted, if necessary, and the iterative procedure continued until a satisfactory result is attained.

IV.—COMPARISON OF RESULTS.

We have now, in effect, by three various methods reduced our original set of measurements with 13 tests to a simpler set of measurements with four factors only. This obviously makes for greater economy in practical descriptions and predictions, quite apart from any concrete 'meaning' the factors may have. But which of the three methods is the best? And what are the differences between the factors they produce?

In general, where practical predictions are required, the least squares method would appear the most trustworthy. As will be seen from the last column of Table II, its factors apparently account for more of the total variance¹ than those obtainable by any other method, and, in particular, its first factor accounts for more than any other first factor. If therefore, we wish to know what will be the probable performances of these children at some future examination requiring all-round educational ability, the least squares method will furnish the best possible estimate, and that in terms of a single factor: if we wish to know a child's performances in each of the separate subjects, it will again furnish the best possible estimate obtainable with a minimum and specified number of factors—one, two, three, or four. The summation method involves rather less labour than the method of least squares, and, as here applied, yields much the same results.

When applied direct to a matrix of variances and covariances² the method of least squares has undoubted advantages. First of all, each person's measurements for the several general factors can be deduced exactly;³ whereas, with the summation method as used by Spearman or Thurstone, the measurements can only be estimated with a certain margin of error. Secondly, from each person's measurements for the n general factors thus deduced we can reconstruct his original measurements exactly; i.e., the method yields an exact fit, not only to the table of covariances or correlations, but also to the list of observed measurements;⁴ whereas from the measurements for the general or common factors obtained by the summation method we can reproduce his observed measurements only approximately (except, of course, when we have obtained estimated measurements for the unique or specific factors too). On the other hand, against the method it is commonly urged that nothing is achieved by converting n tests into the same number

¹ In calculating the percentages in the last column of Table II it is assumed (as usual) that the total amount of variance is equal to the number of tests ($n=13$).

² I have endeavoured to demonstrate the following points by an actual factor-analysis of covariances between temperamental traits. (*Brit. J. Med. Psych.*, XVII, pp. 169 *et seq.*).

³ *Loc. cit.*, p. 175.

⁴ *Loc. cit.*, p. 176.

of factors.¹ The defence, I think, is this. First, there is a definite gain in being able to express the measurements in terms of n *independent* factors instead of n *correlated* tests (and with the least squares method the factor saturations as well as the factor-measurements are uncorrelated); secondly, since each of the factors extracted accounts in turn for a maximum amount of available variance, it is not in point of fact necessary to extract *all* the n factors, but only those few whose contribution to the variance is statistically significant; and, whatever number we extract, we know that with the method of least squares, the squares of the residues will be a minimum for that number.

Thus, with the least squares method, just as with the centroid method, we can (to adopt Thurstone's words) find a "reduced [covariance] matrix of lowest possible rank whose cell entries do not differ from those of the observed matrix by more than might be expected from the sampling errors." And, what is the most important point of all, we can, on taking *any* assigned number of factors (one only, if we like), obtain by this procedure a better fit both to the covariance-table and to the original set of measurements than by any other device.

But, as we have already noted, in dealing with test-measurements expressed in arbitrary units, it is doubtful whether we are justified in working with covariances: and, when we substitute the corresponding correlations, the superiority of the least squares method is by no means so certain. Looking at the figures actually obtained in Table II, it would seem at first glance that the procedure still preserves its main advantages. But a closer scrutiny will show that the general results are in fact very much the same. In practice, the advantage amounts to little more than this: having decided provisionally on a given set of communalities and found saturation coefficients by the summation method, we can always go on by the table-by-column multiplication to get a slightly better fit.²

The group-factor method, on the other hand, does not offer such an exhaustive account of the total variance as either of the foregoing, but it leads to factors that can at once be interpreted. Yet even here a closer inspection will show that there is no real incompatibility between the two results. Certainly, the absolute magnitudes of saturation coefficients, as obtained by the general factor and the group factor methods respectively, differ appreciably; but their relative magnitudes, as we compare

¹ Cf. Thurstone, *Vectors of the Mind*, p. 174. "The solution in which $r=n$ violates the fundamental postulate of science that every valid hypothesis is over-determined." It should be noted, however, that when Thurstone speaks of finding "the smallest number of independent factors that will account for the intercorrelations of n tests," he refers only to common or general factors. On his hypothesis the total number of factors, including specifics, is $r+n$, which is necessarily greater than n , not less. It is indeed just because the total number of factors to be determined is thus greater than the total number of equations for determining them that we can only estimate those unknowns approximately. With the least squares method as described in the above paragraph exactly n factors are required; and hence the number of unknowns is the same as the number of equations (cf. *Brit. J. Psych.*, XXVIII, pp. 76 *et seq.*).

² As noted above, by carrying the process of approximation still further for both methods, the differences between the final results could be reduced almost indefinitely.

one test with another, prove to be very much the same. Thus, if we ask which tests furnish the best measurements of general educational ability, or which are most affected by any particular factor (e.g., the second or verbal factor), the final replies will be identical by all three methods.

This seems to dispose of a common criticism urged against 'general factor' methods. It is, as we have already seen, continually objected that the factors and saturation coefficients they furnish 'can have no psychological meaning.' Even those who adopt the centroid or summation method are nowadays prone to insist that "no attempt can be made to interpret its results directly without a further rotation of axes."¹

In opposition to this view, I have always maintained that the factors obtained by one form of analysis are, almost invariably, related to those obtained by another in some very simple fashion. There are various ways of demonstrating this. In the present case the most direct is to apply a linear transformation to the factor pattern obtained by the general factor methods, so that the factor loadings for that group of tests into which the factor enters positively are maximized and those for the remaining tests (into which the factor does not enter positively) are changed to zero or at any rate minimized. The result should be a close approximation to the pattern obtained by the group factor method. Such a transformation matrix² is shown in Table III.

On applying it to the saturation coefficients obtained by the method of least squares, we reach the saturation coefficients shown in Table IIb. It will be observed that very few of the resulting figures deviate by more than .05 from those deduced by the simple group-factor method described at the outset (Table IIa). If anything, I should prefer the new figures as being more exact: but, in view of the size of the probable error in all

¹ Cf. Thurstone, *loc. cit.*, p. 131, and *Psychometrika*, II, 1937, pp. 74-5. Alexander, *loc. cit.*, pp. 116 *et seq.* Of the method of principal components (or least squares) Thurstone writes: "the reader . . . should not expect to be able to give a psychological meaning to the solution." Holzinger and Stephenson have said much the same about Thurstone's own method. But Thurstone apparently would now always supplement the factor analysis as such by a rotation of axes based on extrinsic considerations.

² The method of rotation most commonly employed by research-students in this country is the graphical method described by Alexander. As a working procedure it has the merits of ease and intelligibility, but (as has already been shown by Aris and others) can only yield very rough results. The method here recommended is to equate the two sets of saturation coefficients and determine suitable multipliers by the method of least squares: when one set of saturation coefficients has already been obtained by the method of least squares the calculations become very simple, since the cross products vanish. The working procedure has already been described more fully by Davies. As will be seen from the figures in Table III, the resulting transformation is not strictly a rotation of axes. It may still be regarded as essentially a change in the frame of reference; but it evidently here involves a slight amount of homogeneous strain as well. In my view, such a transformation is more exact than the rotations usually proposed, and much easier to carry out.

TABLE III.
TRANSFORMATION MATRIX.

<i>Factor.</i>	g_a	v_a	m_a	a_a
g_b	·871	·459	·157	·140
v_b	—·235	·734	—·364	—·333
m_b	·030	—·089	·549	—·473
a_b	—·093	·000	—·138	·534

investigations such as this, there is, for most practical purposes, very little to be gained by first undertaking an elaborate analysis with a general factor method, then determining a transformation matrix by calculations almost as laborious, and finally 'rotating the axes' to obtain results in terms of 'fundamental abilities.' With the group factor method, as must now be obvious, a good approximation can be reached at a single step.

The transformation matrix in itself, however, is not without theoretical interest. It enables us not only to transform the least squares saturations into group-factor saturations, but also to convert the group factor-measurements back into least square factor-measurements.¹ After all, it is the factor-measurements for the individual children that form our ultimate aim: factor analysis, though it looms large in theoretical literature, is only a means, and not always an indispensable means, to that end. As I have argued elsewhere, the psychologist's preoccupation with correlational problems has led him to forget that the operative coefficients are, not the coefficients of correlation, but the coefficients of regression.

¹ This is perhaps not obvious at first sight, but may be proved as follows. Using the notation employed in my previous paper (this *Journal*, VIII, p. 177), let F_a be the 'factorial matrix' of saturation coefficients and P_a the 'population matrix' of inferred factor measurements for each person as obtained by the group factor method ('method a', as it is sometimes called), and let F_b and P_b be the corresponding matrices obtained by the group factor method ('method b'). Let S be the actual scores supplied by the tests, and T the transformation matrix for transforming F_b into F_a . Then $F_a P_a = S$, and $F_b P_b = S$; and $F_a = F_b T$; hence $F_b P_b = F_b T P_a$, and $P_b = T P_a$.

It will be seen from Table III that the matrix T is virtually a 'triangle matrix', i.e., one in which all the elements below the leading diagonal are zero. This, with the inner negative figures, is a regular feature of transformation matrices deriving the group factor saturations from the least squares. The high positive figures in the leading diagonal (decreasing in size in each row) and the smaller negative figures to the right (increasing in size but diminishing in number in each row) are, it will be seen, an inevitable result of what might be called a dichotomous mode of weighting. I have already commented on the emergence of matrices of this type in analysis by the general factor method (cf. 1935 *Memorandum*, p. 307).

Let us suppose, then, that by means of a set of factor saturations like those in Table II_n we have obtained for a batch of children their measurements in what we may call (with Thurstone) their 'primary abilities'—general educational ability, verbal ability, manual ability, and the like; then we can at once re-express them in terms of the least squares factors by taking Table III as a set of weights and recombining them accordingly. These weights, therefore, will express the relations between two kinds of factor. We see at once that the first factor of 'method b' can be identified with the first factor of 'method a' (slightly reduced in relative amount) plus smaller and diminishing amounts of the more specific factors: the second factor of 'method b' virtually expresses the difference between the measurement in the verbal factor of 'method a,' on the one hand, and the measurements in the various non-verbal factors, on the other; similarly, the third factor of 'method b' expresses the difference between the motor factor and the remaining non-motor factors; and the fourth factor of 'method b' is virtually identical with the fourth of 'method a,' i.e., the arithmetical factor, since no non-arithmetical factors now remain to be disposed of.

Thus the first or most general factor with both methods is much the same: but that furnished by 'method b' accounts for a larger amount of the contributed variance (cf. Table II C and D, last column). The special factors are different in their formal character. Those furnished by 'method b' (least squares) are 'bipolar' or 'difference factors': they measure the differences between positive abilities rather than those positive abilities themselves. The special factors furnished by 'method a' (group factor), on the other hand, represent positive abilities only: they add to or increase the effects of the general factor but do not subtract from it or otherwise interfere with it. Thus, 'method b' magnifies the contribution of the first factor at the expense of the secondary factors; 'method a' magnifies the contribution of the secondary or group factors at the expense of the first.

The difference between the two principles is rather like that which traditional logic has always drawn between 'material classification by positive characteristics' and 'formal classification by dichotomous division.' With both methods the first or general factor covers the *totum divisum*, i.e., all the tests or traits selected for experimental study. With 'method a' the remaining factors then simply pick out the verbal, the manual, and the arithmetical processes respectively as co-ordinate groups: with 'method b' the second factor classifies the tests (or rather the processes tested) into verbal and non-verbal; the third then re-classifies the non-verbal into manual and non-manual; and the fourth, had it free play, would evidently sub-classify the non-manual into the arithmetical and non-arithmetical; and so on.

Between the two sets of coefficients, however, there is no real incompatibility: the group factor method builds up a concrete performance by superposed positive capacities; the general factor method first strikes an average, and then expresses further peculiarities as positive or negative deviations above or below that average. The two results are no more 'irreconcilable' than a set of measurements stated first in absolute crude scores and then in terms of positive and negative deviations about the general mean.

V.—STABILITY OF FACTORS.

A second criticism is that the results of a factor-analysis such as the present must be peculiar to the particular set of tests employed and to the particular group of children tested. Beyond question, the validity of any conclusions that can be drawn depends as much upon an appropriate choice of tests and of testees as upon the mode of statistical analysis. But the effects of an unwise or an altered choice must be a matter of degree. Hence we urgently need some measure of factor-stability to discover what difference, if any, such changes have had. Accordingly, to estimate how far a set of factors is constant from one correlation table to another, I have proposed what may be termed a 'symmetry criterion.'

If A and B are any two matrices, then the product AB will not in general be equal to the product BA. But if the factors entering into A and B are identical, then the two products will also be identical; and if A and B, being correlation matrices, are each symmetrical, then it must further follow that the products $AB=BA$ will also be symmetrical.¹

Let us apply this criterion to compare the results obtained with scholastic tests applied to London children nearly 20 years ago with those obtained in the present research. The earlier table of correlations is printed in full in my 1917 *Report*. On multiplying that table by Table I above, we get the product matrix shown in Table IV. Its symmetry is by no means perfect; but the two triangles of coefficients, above and below the leading diagonal, show a reasonably close resemblance. The simplest check is to correlate the totals of the columns with the totals of the rows. The result is a coefficient of .983, where perfect symmetry would, of course, give 1.000.

The chief deviations are displayed by the test of science and (to a less extent) by those of writing, reading, and drawing. But, as we have already seen, science and drawing are the very subjects in which, owing to the changes in teaching and curriculum, my tests had to be considerably revised, while the position of the fundamental subjects of reading and writing has notoriously altered during the last twenty years. It was therefore only to be expected that somewhat different factors should have influenced the children's performance in these particular tests on the two different occasions.

If we confine our comparison to figures from the present research, omitting one or two schools with an unusual curriculum and dividing the children into two comparable groups, the resulting product matrix is almost perfectly symmetrical (correlation .998). If we turn to the test-results for the different age-groups and apply the same criterion, the resulting product-matrices remain almost as symmetrical (average correlation .992), until we get down to the age of 8, comprising children fresh from the very different curriculum

¹ A proof of the criterion has already been given in their theses both by J. M. Woods and by M. Davies, who used it for a similar purpose.

TABLE IV.
SYMMETRY CRITERION : PRODUCT OF TWO CORRELATION MATRICES.

Test.	Composition.	Arithmetic (Problems).	Science.	History.	Geography.	Reading (Comprehension).	Diction.	Handwork.	Reading (Speed).	Writing (Speed).	Arithmetic (Rules).	Drawing.	Writing (Quality).	Total.
Composition.....	3.85	3.33	3.17	3.29	3.24	2.95	2.45	1.95	2.30	2.09	1.87	1.69	1.63	33.81
Arithmetic (Problems) ..	3.06	2.94	2.77	2.64	2.60	2.18	1.90	1.95	1.76	1.79	1.80	1.61	1.62	28.62
Science	2.11	1.80	1.90	1.93	1.82	1.65	1.33	1.14	1.27	1.15	1.00	0.97	0.94	19.01
History.....	2.90	2.57	2.58	2.47	2.43	2.03	1.75	1.69	1.68	1.64	1.46	1.46	1.40	26.11
Geography	2.96	2.59	2.61	2.52	2.51	2.18	1.85	1.72	1.76	1.70	1.52	1.52	1.47	26.91
Reading (Comprehension)	3.04	2.57	2.66	2.60	2.57	2.33	1.93	1.58	1.82	1.67	1.43	1.39	1.33	26.62
Dictation	2.63	2.21	2.31	2.25	2.24	2.02	1.66	1.39	1.59	1.45	1.23	1.24	1.18	23.40
Handwork	2.33	2.15	2.13	2.00	2.01	1.62	1.40	1.71	1.36	1.45	1.36	1.48	1.48	22.48
Reading (Speed).....	2.04	1.78	1.80	1.74	1.73	1.52	1.27	1.16	1.22	1.15	1.03	1.02	0.99	18.45
Writing (Speed)	1.18	0.95	1.01	0.98	1.00	0.85	0.73	0.73	0.72	0.71	0.56	0.71	0.67	10.80
Arithmetic (Rules).....	1.60	1.64	1.47	1.38	1.35	1.09	0.99	1.09	0.89	0.95	1.05	0.87	0.89	15.26
Drawing	2.23	2.11	2.01	1.88	1.90	1.48	1.33	1.67	1.28	1.41	1.36	1.47	1.46	21.59
Writing (Quality)	1.99	1.74	1.80	1.69	1.73	1.40	1.20	1.44	1.19	1.24	1.08	1.29	1.27	19.06
TOTALS	31.92	28.38	28.22	27.37	27.13	23.35	19.79	19.22	18.84	18.40	16.75	16.72	16.33	292.42

of the infants' department. It seems clear, therefore, that the factors emerging from the data here analysed, though subject to minor variations due to differences of age, sex, syllabus, teaching methods, and the like, nevertheless represent at bottom fairly stable components.

For simplicity I have in the foregoing pages accepted the customary language, and referred to these components as 'abilities.' But so far, what we have been examining is really the elementary curriculum, not the mind of the child. If our intention were to classify or grade the individual children, we should now proceed to draw up a series of regression equations, and then compute the various factor-measurements in detail. In that way we might hope to study the distribution of the different 'abilities' and 'types' which the analysis seems to suggest. So far, however, we have only classified school subjects, not school pupils.

This will perhaps become clearer if, once again, we think for a moment of correlating, not tests or school subjects, but persons. Such a procedure would be perfectly legitimate here, because the subjects, though in no sense a normally distributed sample, nevertheless form a limited and well-defined group: we thus escape the troublesome problems that arise in other fields, where we correlate between, or over, a highly arbitrary collection of tests or traits. The saturation coefficients for persons (which will be similar to, if not actually identical with, the factor-measurements for persons obtained by correlating tests) would then evidently indicate the 'type' to which each child approximates and the degree of his approximation: but the factor-measurements for tests (which will be similar to, and in certain cases identical with, the saturation coefficients for tests obtained by correlating tests) must evidently indicate the nature of the educational subjects tested.

The fact is that, whichever line of approach we follow, what we analyze is not the child *in vacuo*, but the *relations* between children and their educational environment. In my view the notion of mental 'abilities' should now be accorded the same status in psychology as the notion of physical 'causes' in material science: both may be convenient terms for popular exposition, but both are highly inconvenient for scientific thinking. In physics, where the plain man pictures a physical cause residing in one substance and operating on another, there the physicist will merely talk of 'functions'; and his 'functions' will nowadays be expressed by the 'characteristic values' of 'linear operators' (or 'matrices'): in short he will deal with systematic sets of relations, and never with the terms that are related except so far as those terms are themselves reducible to systems of relations. The same is true of the factors reached by mathematical analysis in psychology. This caution is appended here because it has more than a mere metaphysical bearing. In the past the medical psychologist, dealing with a youngster whose work or behaviour was subnormal, would test or examine the child in his clinic or consulting room, and think his task was ended when he had measured the mental abilities of the child or assessed his innate propensities. A sounder practice is coming to realize that nearly always the 'problem' lies, not exclusively in the 'problem-child,' but in the complex relations between the child on the one hand and his environment on the other. In the last resort, therefore, our factors are abstract aspects of a reciprocal interaction, not concrete entities lodged in something that is loosely called the mind.

VI.—SUMMARY.

(1) The object of the present inquiry was to re-investigate the relations between performances in the chief branches of the school curriculum by means of more modern statistical methods, and to determine how far those performances could be explained by a single general factor, and whether they implied (as was originally suggested in an earlier research) a number of more specialized 'group factors.'

(2) 613 ten-year-old children have been tested by standardized tests of the usual elementary subjects, and the correlations submitted to factor-analysis both by a 'group factor method' similar to that employed in the earlier research and by more recent 'general factor methods.' With the increased numbers in the sample the probable errors are now so small that no doubt can remain about the statistical significance of the secondary or 'group' factors emerging. Whichever procedure is used, the factors reached are much the same as those originally described: namely, (i) a general, (ii) a verbal, (iii) an arithmetical, and (iv) a manual factor.

(3) The group factor method proves to be a direct and easy way of isolating a 'simple structure' of 'primary abilities,' i.e., a factorial matrix in which there shall be a maximum of zero coefficients and a minimum of negative coefficients. A triangular transformation matrix is derived which enables the results obtained by 'general factor methods' to be interpreted in terms of the primary factors furnished by the 'group factor method': the essential relation is that the factor-measurements obtained by the former methods express the weighted differences between the factor-measurements obtained in terms of 'primary abilities.'

(4) On applying the 'symmetry criterion' it is shown that within schools of the same type the factors appear fairly stable from one batch of pupils to another: they thus serve to account for the chief recurrent types of general and specialized ability and disability popularly recognized among children as judged by their work in the elementary school.

RÉSUMÉ.

LES RAPPORTS DES APTITUDES SCOLAIRES.

Le but de l'enquête actuelle était de décider si les rapports entre les aptitudes scolaires pouvaient s'expliquer uniquement par un seul facteur central ou si (comme l'on suggéra dans une recherche antérieure) ils impliquent aussi l'existence de facteurs collectifs. 613 enfants de dix ans ont été examinés au moyen des tests établis dans les domaines de l'instruction les plus répandues, et les corrélations soumises à l'analyse en facteurs par chacune des méthodes bien-connues. En

particulier la "méthode des facteurs collectifs," employée d'abord en 1913, a été rappliquée, et l'on a découvert qu'elle fournit des résultats qui concordent avec ceux donnés par des méthodes plus récentes; à vrai dire, elle semble offrir un moyen beaucoup plus facile et plus exact de déterminer "les aptitudes intellectuelles primaires" qui ceux qui obligent à dresser d'abord une série de facteurs indépendants au moyen de quelque "méthode de facteurs centraux," pour faire "tourner" ces axes de facteurs jusqu'à ce qu'on obtienne une "structure simple." Les mesures des facteurs fournies par les méthodes de facteurs centraux expriment, comme on le démontre, les différences entre les mesures des facteurs fournies par les méthodes des facteurs collectifs.

Les facteurs découvertes à la fin sont presque identiques avec ceux de la recherche originale, à savoir, (i) un facteur central, (ii) un verbal, (iii) un arithmétique et (iv) un manuel. Le "critère de la symétrie" démontre que les facteurs restent à peu près constants chez les différents groupes d'élèves; ils semblent donc expliquer les formes principales de l'aptitude et de l'inaptitude, générale et spécialisée, qu'on peut remarquer chez les élèves.

ZUSAMMENFASSUNG.

DIE VERHÄLTNISSE ZWISCHEN ERZIEHERISCHEN FÄHIGKEITEN.

Die vorliegende Untersuchung sollte feststellen, ob die Verhältnisse zwischen erzieherischen Fähigkeiten allein durch einen einzelnen allgemeinen Faktor erklärt werden können, oder ob (wie anfangs in einer früheren Untersuchung angedeutet wurde) sie auch auf das Vorhandensein von Gruppenfaktoren hinwiesen. 613 10-jährige Kinder wurden durch Standardtests der üblicheren Schulfächer geprüft und die Korrelationen wurden der Faktorenanalyse durch jede bekannte Methode unterworfen. Insbesondere hat man die ursprünglich 1913 gebrauchte "Gruppenaktorenmethode" von neuem angewendet. Diese hat Resultate ergeben, die mit den Resultaten neuerer Methoden übereinstimmen: sie scheint in der Tat ein viel leichteres und zuverlässigeres Mittel zu sein, um geistige Grundfähigkeiten zu bestimmen als jene bei denen zuerst eine Gruppe unabhängiger Faktoren durch eine "allgemeine Faktorenmethode" hergeleitet werden muss, sodass man diese Faktorenachse "rotieren" lässt, bis "ein einfacher Bau" erzielt wird. Die von den allgemeinen Faktorenmethoden gelieferten Faktorenmassen weisen Unterschiede gegenüber den von Gruppenfaktorenmethoden gelieferten Faktorenmassen auf.

Die schliesslich entdeckten Faktoren sind mehr oder minder identisch mit denen, die man bei der ursprünglichen Untersuchung entdeckt, nämlich:

- (1) Ein allgemeiner,
- (2) Ein Wort-,
- (3) Ein Rechen-, und
- (4) Ein Fingerfertigkeitstfaktor.

Das "Kriterium der Symmetrie" zeigt, dass die Faktoren von einer Schülergruppe zur anderen wesentlich dieselben bleiben: sie scheinen also für die Hauptformen allgemeiner und besonderer Fähigkeit und Unfähigkeit, die bei Schulkindern zu beobachten sind, Rechenschaft abzulegen.

GENERAL KNOWLEDGE AND INTELLIGENCE.*

By DOROTHEA M. INMAN.

- I.—*Purpose and scope of the experiments.*
- II.—*Study of the correlations observed.*
- III.—*Case-studies: the affective element.*
- IV.—*Differences due to sex and environment.*
- V.—*Summary of results and conclusions.*

I.—PURPOSE AND SCOPE OF THE EXPERIMENTS.

ALTHOUGH in many schools a general knowledge paper is an annual event, very little research has yet been done on the subject, and in this country nothing seems yet to have been published on the correlation between the results of general knowledge papers and intelligence tests.

The research here described was, unfortunately, much restricted by lack of time and funds for its extension, but the results obtained were in sufficiently close agreement with those of scholars working in comparable fields to suggest that further research on a larger scale would lead to valuable and interesting conclusions.

The subjects in this enquiry were 94 boys and 107 girls, all between the ages of twelve and fifteen and a half years. They were drawn from two mixed secondary schools, one in an eastern and one in a western suburb of London (hereafter referred to as E.S.S. and W.S.S. respectively), and from a High School of the Girls' Public Day School Trust and a boys' preparatory school in the same wealthy residential district. The average ages of the groups were as follows:

E.S.S.	Boys, 14 years 4 months;	girls,
	14 years 5 months.	
High School	Girls, 13 years 9 months.	
Preparatory School ..	Boys, 13 years 0·8 months.	
W.S.S.	Older boys and girls, 14 years;	younger
	boys and girls, 13 years.	

The writer is very much indebted to the head masters and head mistresses of these schools for much valuable assistance.

The children were given two parallel general knowledge papers, each comprising questions on school interests and out-of-school interests.

*A summary of a thesis, entitled "A Study of the General Knowledge of London School Children in relation to their Intelligence," approved in December, 1936, for the degree of Ph.D. of London University. Available in the Library of the University of London.

They were also given the Otis Group Intelligence Scale (Advanced Form A), except that in one school where the children's intelligence quotients were already known only a small sample took this test. As the Otis test was found to yield an appreciably higher I.Q.¹ the results of the previous test were scaled up for comparison with the other groups.

In order to free the general knowledge papers, as far as possible, from bias in favour of any one type of child, they were first of all divided into two main sections: school interests and out-of-school interests. These were sub-divided as follows:

<i>School Interests.</i>	<i>Out-of-School Interests.</i>
(1) Mathematics; Weights and Measures.	(1) Mechanical Interests.
(2) Scientific Interests.	(2) Political and Civic Knowledge.
(3) Nature Study.	(3) Light Literature; Children's Books.
(4) Hygiene and Domestic Science.	(4) Stage and Screen.
(5) Geography.	(5) Proverbs, Fairy-Stories, and Nursery Rhymes.
(6) History.	(6) Knowledge of London.
(7) Standard Literature.	(7) Abbreviations and Foreign Phrases.
(8) Art.	(8) Trades and Professions.
(9) Music.	(9) Famous People of To-day.
(10) Biblical and Mythological Allusions.	(10) Sport.

Each of the sub-sections contained ten questions, giving a total of 200 questions in each of the two general knowledge papers A and B. An attempt was made to pair actual questions according to difficulty, but this was only moderately successful. The nature of the subject-matter made it impossible for each sub-section to be anything approaching a "pure test" of knowledge in the field from which it took its name. A fact best claims to rank as "general knowledge" when it affects life at many points. For this reason most of the questions dealt with superficial aspects of the subject-matter, and answers showing only superficial knowledge were marked as correct.

Because of the "impurity" of the tests and the small number of questions in each sub-section the results of the A and B forms of these sub-sections taken individually gave only low coefficients of reliability, mostly in the region of .50 and .60, but falling once as low as .08. On the other hand, the reliability coefficients obtained by comparing the results of the two whole papers varied with different groups from

¹ Cf. E. G. BRADFORD on "Birth-rate and Intelligence," commenting on London results obtained by the Otis Test.—*This Journal*, Vol. VII, Nov. 1937, p. 233.

·99 to ·84, and those obtained from the two main sections from ·98 to ·79. While, therefore, the sub-sections gave little guidance to the subjective interests of individual subjects, it is probable that the whole tests gave a fair sample of their mental background.

The questions were so framed as to require short answers of the reproductive type. This type of question was preferred to the true-false form or the multiple choice response because it was felt that the two latter forms of question admit the exercise of intelligence to a greater degree than the former, and that their use would, therefore, vitiate the comparison with the results of intelligence tests. Many sub-sections would have been more valuable if it had been possible to illustrate the tests, but consideration of expense prevented this. Partly because of this the tests probably favoured the child with a biographical interest in such subjects as art and music.

It was originally intended that all the sub-sections should be of equal difficulty. In spite of considerable changes made after the preliminary experiments and of lesser ones made after visiting the second school this equality was not attained. The last school visited was the only one in which children of two successive school years were tested: in this school the mean score of the older group showed a superiority to that of the younger which was proportionately greatest in those sections (standard literature, art, music), which had been found most difficult by all subjects. It is therefore probable that had the age range been extended upwards the profile obtainable from the mean scores in each sub-section would have tended to smooth out, while the examination of a group of adult subjects would have afforded material for a most interesting comparison.

II.—STUDY OF THE CORRELATIONS OBSERVED.

From each school tested some rating was obtained for school work. In two schools this took the form of numerical marks, in one of ranks, and in the fourth of letter-grades within which there was no ranking. In each group the correlation was found between general knowledge (total score) and mental age, between general knowledge and school work, and between mental age and school work. For reasons which need not here be enumerated the correlations with school work had not the same statistical value as those between general knowledge and mental age, but they were suggestive of interesting possibilities.

For all groups taking the final draft of the general knowledge paper the weighted average correlation between mental age and general knowledge was ·45, between school work and general knowledge ·41, and

between mental age and school work .27. In the two schools first visited, where the general knowledge papers differed slightly both from each other and from the final draft the correlations were as follows :

	<i>G.K. and M.A.</i>	<i>G.K. and S.W.</i>	<i>M.A. and S.W.</i>
E.S.S. :			
Boys56	.32	.47
Girls34	.33	.32
High School66	.57	.38

In almost every group the correlation was highest between general knowledge and mental age, and lowest between school work and mental age. None of the differences was of statistical significance, but these figures agreed closely with the conclusions of Terman and his colleagues.¹ Among the tests given to the "gifted children" was one of general information, and Terman writes of the results: "That the gifted make a showing in General Information even better than they make in such subjects as reading, arithmetic, and spelling, is probably due to the fact that the child's stock of information is more dependent upon intellectual initiative and less upon formal school instruction. It would seem, therefore, that general information tests might be more valid than achievement tests for selecting pupils for instruction in gifted classes. They probably compare well with the best group intelligence tests for this purpose, and they have the advantage of being easier to administer."

The striking thing about the figures obtained by the writer is the lowness of the correlations between general knowledge, mental age, and school work. One reason for this is that the groups tested were small; the largest group for which the three figures required for comparison could be obtained was one of twenty-nine children. In addition, the subjects were drawn from a highly selected portion of the population, all the secondary school children having been special place winners, while those from the High School and preparatory school came from homes much above the average national income level. The I.Q.'s found by the Otis Test ranged from 160 to 98; the mean I.Q. was not calculated for the whole number of subjects, but the mean for the lowest group was 121. While, as observed earlier, this figure was probably too high, the subjects were undoubtedly well above average in intelligence.

¹ TERMAN, L. M., *et al.*: *Genetic Studies of Genius*, Vol. I, p. 304. (Stanford University Press, 1925.)

Terman (op. cit., p. 582) quotes a correlation " between .80 and .90 with the Terman Group Test " for his General Information Test, a figure apparently based on the experiment with the high school students. This much higher figure is probably due in part to the facts that the Terman Group Test itself contains a test of information, and that the Terman Information Test was in the multiple choice response form deliberately avoided in the writer's experiments.

While the writer's experiments show a substantial correlation between general knowledge and intelligence, it is clear that with small groups of highly selected subjects the influence of other factors must considerably affect the coefficients. The next section, on case-studies, indicates what some of these factors may be.

III.—CASE-STUDIES : THE AFFECTIVE ELEMENT.

Since the coefficients of correlation obtained between general knowledge and intelligence, though reduced by the factors mentioned above, remained sufficiently substantial to support the common sense hypothesis that the more intelligent a child is, the greater, other things being equal, will be his stock of general knowledge, individual study was made of some twenty children whose scores in the two tests were widely discrepant. All the children tested had supplied particulars of the occupation of their parent or guardian, of the number of children in the family and their own position (eldest, youngest, etc.) within it and of their exact age. These particulars were taken into account in selecting cases for special study. In one school only the writer was told before selecting cases that a certain boy was regarded as a problem. He was later chosen for case-study on the strength of his scores and not of this observation.

Each child studied was interviewed for about twenty minutes, in which he or she was asked questions about the home-life, including the time taken up by household duties, interests in and out of school, tastes in films and wireless programmes, and any sections of the general knowledge paper in which the child's score was remarkable. These interviews with the children were supplemented by interviews with the head teachers and the class teachers, and in one school by a study of school reports.

Two cases will be described, one that of a girl with very high marks in general knowledge and relatively low ones in the intelligence test, the other that of a boy who reversed these positions :

(1) *May C.* was tested and interviewed at the first school visited. Age 15 years 2 months. Average age of class, 14 years 5 months. Father, collector and salesman. One elder brother.

*Standard Scores*¹ (based on mean and S.D. of the girls in this school) :

General Knowledge Total	+2.33
Otis Score	-0.42
I.Q.	-0.97
School Work	+0.70
Age	+1.46

Notable Sections in G.K. :

Geography	+2.98	top of both boys and girls.
Standard Literature	+2.18	Ditto.
Music	+3.65	(this test showed very great variation in all groups tested).
Stage and Screen	+2.93	Second of boys and girls.
Proverbs, etc.	-2.20	

The discrepancy between the scores for the Otis Test and general knowledge is very large, but there seems no reason to suppose that the former had done May any serious injustice. The head master described her as a plodder. It was clear that she came from a very poor home, and that her health was indifferent, probably owing to unsuitable food and inadequate sleep. The only obvious reason for her superiority in general knowledge was that she was seven months over the average age of the class.

When interviewed, May proved friendly, not at all shy, and decided in her views. Little of her time was taken up by travelling or household duties. She took a keen interest in political news, both in the newspapers and on the wireless. Her favourite reading consisted of the better type of detective story and a popular film weekly. She went to the cinema once a week and preferred historical and humorous films. Other hobbies were knitting and painting, and at another school she had belonged to an art club. Her favourite school subject was History, and, in spite of her high score, she denied any special interest in Geography. She indignantly denied any interest in classical music, but attributed her unusual knowledge of composers' names to the wireless, explaining: "You hear the announcer give out who it's by, and then you turn it off."

In only three sub-sections did her score fall below the average, these being Nature Study, Hygiene and Proverbs, Fairy-Stories and Nursery Rhymes. Weakness in the first was probably due to the fact that May had never been into the country, while the poverty of the home probably accounted for the second. The class-mistress made the interesting suggestion that poverty might also cause the weakness in the last of these three sections, the mother being too busy to amuse the child in this way.

¹ The Standard Score is obtained by expressing the individual score in terms of its difference from the mean of the group, and then dividing by the S.D. This makes possible accurate comparison of the results of different individuals in the same test or of the same individual in different tests.

May had no elder sister, nor was there any younger child whom she herself might have had to amuse.

Taking these various circumstances into consideration, it seemed that May's fund of general knowledge was due to her own personality and general attitude to life, together with her chronological age, which, as noted, was above the class average. Her intelligence was good, though below the average of a highly selected group. She showed generally sound tastes and an unusually lively interest in the world outside school and home. Not only was she above the average chronological age of the group, but she gave the impression of unusual maturity for her fifteen years. It might be said that she knew a great deal because she had made up her mind to do so.

(2) *Kenneth G.* was tested and interviewed in the last school visited. Only child. Mother factory-hand. Father deserted mother five years before interview. Age 12 years 7 months. Average age of class 13 years. *Standard Scores* (based on mean and S.D. of boys in own age-group in this school) :

Age.....	—0·95
Mental Age	+0·86
I.Q.	+1·50
School Work	—1·33
G.K. Total	—0·42
School Interests	—0·42
Out-of-School Interests ..	—1·03

Kenneth's score, with a marked difference in his performance in the sections on school interests and out-of-school interests, together with the absence on his list of routine particulars of any reference to his father, suggested that a restricted environment might account for much of the weakness in general knowledge, and this proved to be the case. At the same time, it soon became clear that this weakness was not due to poverty alone, but also to defects of character aggravated by poverty and an unhappy home. It will be noted that Kenneth's standard score in school work was considerably lower than that gained in general knowledge. Members of the school staff were sceptical of the accuracy of the intelligence test, but the head master regarded Kenneth as a clever boy presenting an acute problem. His school reports commented on his lack of concentration, while his conduct had been unsatisfactory from his first entry into the school. Shortly before the interview he had been detected in the theft of two stamp-albums, which, after two days' truancy, he had returned by post with an anonymous note purporting to come from the mother of a penitent and severely punished offender. The recency of this episode no doubt increased the nervousness shown during the interview.

Kenneth's interests and hobbies seemed normal for a boy of his age.

His newspaper reading consisted of the sports page and serial, with an occasional glance at the political news. On the wireless he preferred plays and the news-bulletin, but as he and his mother "lived with friends" his freedom of choice was restricted. His leisure reading consisted of "2d. bloods," Zane Grey and Edgar Wallace. While at his junior school he had borrowed from the public library, but no longer did so, this being possibly a symptom of increasing maladjustment. Until about three months before the test he had hardly been to a cinema, but had recently been about once a fortnight and had developed a taste for "thriller" films. His favourite school subject was science, and his hobby making scientific models and collecting stamps—legitimately as well as otherwise. He wished to become an electrician. In spite of his scientific interest he had done conspicuously badly in this section.

While Kenneth was somewhat below the average age of the class, this alone would not account for the lack of General Knowledge. Poverty, with the accompanying restriction of environment, was undoubtedly a direct cause, while to a boy the absence of a father would be an especially great disadvantage. In this case it seems probable that before the father deserted his wife the home had been an unhappy one and that the emotional strain had hampered Kenneth's psychological development.

Of the twenty completed case-studies, ten were of children whose general knowledge was surprisingly good as compared with their Otis scores, and ten of children whose general knowledge was, by the same standard, surprisingly poor. In each group of ten five cases were below the average age for their class and five were above it. The greatest variations from the average age were found among the children with good general knowledge. The case of May C. has been described above; Lily B., another girl in the same class, was 14 months (2.29 S.D.) above the average age. Lily had been kept back in school work by an illness which had not prevented her from reading extensively. In the W.S.S. one of the boys interviewed, Tony M., was six months (1.14 S.D.) above the average age of 13 years.

On the other hand, in the same group of children with good scores in general knowledge were the following cases:

<i>Name.</i>	<i>School.</i>	<i>Age.</i>	<i>Average Age.</i>	<i>Standard Score.</i>
Ian L.	E.S.S.	13 yrs. 8 mths.	14 yrs. 4 mths.	-1.80
Denis B.	Preparatory	12 yrs. 8 mths.	13 yrs.	-1.12
Rosalind W.	W.S.S.	12 yrs. 6 mths.	13 yrs.	-1.42

Of these three children, Denis B. was so exceptionally tall and physically developed that his head master was under the impression that he was a year older than his actual age ; Rosalind W. was a Jewess, and also somewhat advanced in physical development ; Ian L., who came from a more educated home than most of his school fellows, devoted to scientific hobbies, reading, and listening to wireless talks more time than most boys of his age, because his sight kept him from much active part in sport.

No child studied because of his or her low score in general knowledge differed from the average age of the class by as much as 1 S.D.

The figures obtained lead to no definite conclusion as to the part played by chronological age in determining the growth of general knowledge. This was natural, since the scores in each case were compared with the mean of a single school class, in which the actual age range was always small. Moreover, within groups so selected, the elder children tend to be the less intelligent. The coefficient of correlation between General Knowledge and chronological age was calculated for the E.S.S., but was found, in the circumstances, to be insignificant, and the calculation was not repeated for the later groups.

The two cases described in full were the most striking met, but the rest supported the conclusion that those children whose score in general knowledge was markedly greater than that gained in the intelligence test were in some way unusually mature either for their actual age or at least for the class in which they worked. All seemed, on the evidence available, to be well-adjusted children. Those, on the other hand, whose score was markedly lower in general knowledge than in the intelligence test, seemed, on the same evidence, to be somewhat childish in their general attitude, being for different reasons too much absorbed in their personal problems for their interests to have widened to the same extent as those of their contemporaries. One girl's deficiency in general knowledge was largely attributable to a foreign home background coupled with the comparatively recent death of her mother, while one boy suffered from a severe sight-defect, with consequent restriction of reading, but both these children showed some signs of maladjustment, the girl's probably temporary. Of the nine cases studied because of lack of general knowledge, three were children from broken homes, while the home of a fourth was known to be unsatisfactory.

Although the number of cases examined was insufficient to support any dogmatic conclusion, it seems probable that a test of general

knowledge given with a test of intelligence might serve as a rough test of emotional maturity. A. G. Hughes in an article on "Discrepancies between the Results of Intelligence Tests and Entrance Examinations to Secondary Schools"¹ summarises as follows the conditions necessary for success in secondary school work :

- " (1) a good degree of innate intelligence ;
- (2) a certain standard of academic attainments on entry ;
- (3) some sturdiness of character, preferably combined with physical vigour and satisfactory home conditions."

This corresponds closely with Professor Valentine's findings recently set out in "Examinations and the Examinee." The experiments here described suggest that the same conditions, or at least the first and third, are necessary for success in a test of general knowledge, although the amount of non-scholastic material included gave an opportunity to children less well adjusted to school work than to life outside. One or two children whose school marks were low for their score in the intelligence test did well in the general knowledge papers.

IV.—DIFFERENCES DUE TO SEX AND ENVIRONMENT.

(a) *Sex.*

In both the secondary schools tested the average score of the boys in the general knowledge paper was considerably higher than that of the girls of the same chronological age. The difference was found both in the School Interests and in the Out-of-School Interests sections, and in all sub-sections except those on Hygiene and Domestic Science and on Proverbs, Fairy Stories, and Nursery Rhymes. It was greatest in those on Mathematics, Science, Geography, Mechanical Interests, and Sport, and least in those on Art, Music, and Literature (both standard and light). It was also, in view of tradition, rather surprisingly small in the section on Political and Civic Knowledge, perhaps owing to the teaching of "civics" in some form in the schools.

Very similar differences were found between the results of the High School girls and the preparatory school boys, but the two groups could not be fairly compared since the boys were younger, but much more intelligent and a large number of them came from homes of high professional standing.

¹ This *Journal*, Vol. IV, November, 1934.

In comparing these results it is necessary to make some allowance for Intelligence. The following table sets out the differences observed :

	G.K. Total.		Dif- ference	P.E.	Otis Score.		Dif- ference	P.E.
	Boys.	Girls.			Boys.	Girls.		
E.S.S.	191.00	148.76	42.24	5.83	181.36	174.72	6.64	2.60
W.S.S.								
Elder	159.83	116.50	43.33	6.59	164.35	153.42	10.93	3.48
Younger	123.30	98.25	27.05	4.93	150.93	147.07	4.86	3.29
All	141.52	106.67	34.85	5.09	156.87	150.26	6.61	2.69

The actual ages, quoted in Section I, were approximately equal for boys and girls.

The above figures suggest that quite probably some factor other than chance must be held responsible for the lower intelligence of the girls among the subjects tested, but it would be beyond the scope of this enquiry to speculate what this might be. A comparison of the differences between the means for the Otis Test and for general knowledge, each compared with its own Probable Error, strongly reinforces the conclusion that the difference in intelligence between the boys and girls tested can account only for a small part of the difference in general knowledge. This, it will be seen, was found both in the E.S.S., where boys and girls were taught separately and had slightly different curricula, and where the girls had rather higher marks for school work, and in the W.S.S., where the two sexes were taught side by side, but where the girls had a lower mean score for school work. Terman (op. cit., pp. 298 ff) found boys superior to girls in General Information, both in the gifted and in the control group. The differences between the forms of his tests and those described make more detailed comparison difficult, but the general agreement suggests the existence of a real sex-difference needing explanation.

The most obvious reason for the inferiority of girls in general knowledge is in the difference in the out-of-school life of the two sexes. Although no girl interviewed seemed overburdened by domestic duties, there are very few girls of the social classes from which these subjects were drawn who do not spend more of their out-of-school time in household tasks than their brothers do. Feminine hobbies such as needlework and in most cases knitting further curtail reading-time. Boys still enjoy more liberty than girls in the use of their leisure and spend more of it outside the home.

On the other hand, Burt and Moore, in their well-known study of "The Mental Differences between the Sexes"¹ found that "Feminine superiority is a constant phenomenon in memory tests of every kind." Common sense suggests that general knowledge depends largely on memory, and it seems strange that the sex found experimentally to be stronger in memory should be inferior in general knowledge. Freud and other psychologists before him have taught us the importance of motive in memory, and if we tend to forget the disagreeable we also remember the useful. In laboratory tests it may be assumed that all subjects are motivated alike: hence the tests measure pure capacity. In school and in life, on the other hand, motivation varies between individuals and between the sexes in respect of different material. The boys showed most superiority in those sections of the General Knowledge paper which tested knowledge more likely to be of use to them than to girls in after-school life, while the two sections in which the girls excelled followed the same rule. Girls again, although most of them now follow a paid occupation for some years after leaving school, have, since most of them abandon this work on marriage, on the whole less motivation than boys to improve their social and financial status by their professional efforts, and so, perhaps, also less motive to acquire and retain general knowledge in the widest sense. In the section on Trades and Professions some questions dealt with careers open mainly to men and some with those more likely to be chosen by women. The percentages of correct answers given to these questions by girls and boys respectively supported the view that a vocational interest is likely to influence the acquisition of knowledge.

It will be seen that the differences observed in different sections bear some correspondence to the views commonly held among teachers on sex-differences in various school subjects. Those in which boys are traditionally superior are those in which their vocational interest is stronger than that of girls.

Making all allowance for vocational interest and for environmental differences, a detailed study of the questions in which boys showed most superiority still left some evidence in support of the common view that boys show more initiative than girls. It is possible that this is especially apparent within the age-range studied, since during the earlier years of adolescence girls are more liable than boys to feel strain due to the changes attendant on puberty.

The effect of chronological age on this sex-difference in general knowledge could only be traced in the W.S.S., where children of two successive school years were tested, and where the average ages of the

¹ *Journal of Experimental Pedagogy*, Vol. I, No. 5, December, 1912, pp. 325-388.

boys and girls in each group were equal to within a fraction of a month. Here the mean score of the younger boys was superior to that of the elder girls in the general knowledge total, in both main sub-sections, the difference being greater in the section on Out-of-School Interests, and in eleven of the twenty detailed sub-sections. The following table shows a curious difference between the sexes in the margins whereby the older classes surpassed the younger. The first column shows the percentage by which the mean of the older group exceeds the mean of the younger, while the second shows the difference between these means in terms of its own Standard Deviation.

	Per cent Increase.		Difference of Means ÷ S.D. Difference of Means.	
	Boys.	Girls.	Boys.	Girls.
Whole Papers	29·60	21·10	3·37	3·03
School Interests Section	27·60	28·30	2·50	3·20
Out-of-School ditto	32·00	12·70	3·53	1·72

While in proportion to their lower mean the older girls showed in the School Interests section a slightly greater superiority to the younger girls than did the older to the younger boys, in the section on Out-of-School Interests they were superior only by a margin which might easily have occurred by chance. The older boys, on the contrary, were superior to the younger boys by a much larger margin in this section than in that on School Interests.

This difference possibly reflects an increase in the domestic duties of the older girls, with a consequent restriction of leisure, and possibly also the incidence of the strain of puberty. The average age of the older boys and girls was in this school fourteen years, and of the younger thirteen. It so happened that the average I.Q. of both boys and girls was higher in the younger than the older groups, the difference in the means being 3·92 points for the boys and 4·93 points for the girls. The difference was therefore slightly greater between the two groups of girls, and presumably accounted for some part at least of the difference between the margins of superiority in favour of the older group. It would not affect the difference noted between the advances in the two main sections of the papers.

It will be observed that the superiority of the boys in general knowledge was slightly greater in the fourteen year old group than in the thirteen year old.

(b) Environment.

The number of subjects tested did not give adequate material for the study of the effect of environment. The High School girls came from homes on the average wealthier than those of the secondary school girls, but they were less intelligent. Many of the preparatory school boys came from exceptionally cultivated homes, but their average I.Q. was eight points above that of the boys of the same age at the W.S.S., so that this must account for a large part of their actual superiority in general knowledge. The superiority was greatest, as might have been expected, in the two sections on literature and in those on Art and Music, and greatest of all in Geography, a subject which in that school seemed to be taught in a particularly stimulating way. Some more specific differences between the wealthier and less wealthy children could be found, for example, in the range of sports celebrities with whom they were familiar, "Rugger" being better known by the preparatory school boys and dirt-track-racing to those from the secondary schools. While some of the individual case-studies supported the view that poverty produces backwardness in general knowledge as well as in school work, the tests covered a sufficiently wide range of subjects to allow for considerable variations in interest without great loss in the total score. On the whole the school curriculum seemed to influence success in different sections of the test more than did home circumstances. The influence of current events was clearly traceable in the fluctuations in percentages of correct answers to some questions. The tests were spread out over eight months ranging from just after the Silver Jubilee until just after the death of King George V. It is hardly surprising that within that period the children showed a remarkable familiarity with all questions touching the Royal Family in any way, but in many other questions there was striking uniformity in the percentage of correct answers. This was greatest in the section on Famous People of To-day. From a list of twenty assorted celebrities which included Franklin Roosevelt and Jean Batten, Mrs. Dionne and Sir John Reith, 98.5 of all the children could give an adequate account of Henry Hall and 97.5 of Walt Disney. Mr. Roosevelt came third on the Roll of Fame.

The influence of broadcasting and the cinema was clearly traceable, the latter making it difficult to accept the results of the section on Light Literature as any indication of what had actually been read.

V.—SUMMARY OF RESULTS AND CONCLUSIONS.

The subjects in this experiment, 201 boys and girls between the ages of twelve and fifteen and a half years, drawn from four different schools, were given tests in general knowledge, the results of which were correlated with those of a group intelligence test, and also with school marks.

The mean correlations observed were .45 between general knowledge and mental age, .41 between general knowledge and school marks, .27 between mental age and school marks. These figures support the conclusion reached in previous American studies that there is a substantial correlation between tests of intelligence and of general knowledge. The relatively low coefficients of correlation obtained in this experiment was partly attributable to the restricted selection of subjects and partly to differences between these tests and those used in the American studies.

Case-studies suggested that temperamental factors were also involved, and that the ill-adjusted child was likely to be slow in acquiring general knowledge.

Throughout the age-range studied boys were found to be much superior in general knowledge to girls of the same age. The difference may be in part accounted for by a difference in out-of-school environment and in part by a stronger vocational motive among boys to acquire and retain information on a wide variety of subjects.

It was impossible to reach definite conclusions as to the influence of social environment.

RÉSUMÉ.

LE RAPPORT ENTRE LES CONNAISSANCES GÉNÉRALES ET L'INTELLIGENCE.

L'on fit passer deux examens parallèles sur les connaissances générales à 201 garçons et filles, élèves de deux écoles secondaires mixtes de la banlieue de Londres, et d'un lycée de jeunes filles et d'une école préparatoire de garçons dans un quartier où demeurent des gens aisés. L'on établit les corrélations entre les résultats de cet examen, ceux d'un test collectif et les notes scolaires. La corrélation moyenne remarquée entre l'âge intellectuel et les connaissances générales était de .45, entre les connaissances générales et les travaux scolaires de .41 et entre l'âge intellectuel et les travaux scolaires de .27.

Une étude de cas individuels suggère que le niveau de développement social était un facteur dans l'acquisition des connaissances générales dans lesquelles l'enfant mal-ajusté se montrerait probablement défectueux.

L'on trouva les garçons supérieurs aux filles dans tous les groupes d'âge étudiés ; cette différence peut s'attribuer en partie à des différences dans le milieu en dehors de l'école et dans l'attitude vers le choix d'une profession.

ZUSAMMENFASSUNG.

DAS VERHÄLTNIS ZWISCHEN ALLGEMEINEN KENNTNISSEN UND DER INTELLIGENZ.

201 Knaben und Mädchen zwischen zwölf und fünfzehneinhalb Jahren aus zwei koedukationellen höheren Schulen in den Vororten Londons, aus einem Mädchenlyzeum, und aus einer privaten Knabenvorschule in einem wohlhabenden Wohnviertel unterzogen sich einer Prüfung in allgemeinen Kenntnissen. Die Ergebnisse wurden mit einem Gruppenintelligenztest und mit den Schulleistungen verglichen. Die beobachtete Mittelkorrelation zwischen geistigem Alter und Kenntnissen ergab sich als ,45, zwischen Kenntnissen und Schulleistung ,41, und zwischen geistigem Alter und Schulleistung ,27.

Falluntersuchungen wiesen darauf hin, dass gesellschaftliche Reife einen Faktor in der Erwerbung umfangreicher allgemeiner Kenntnisse bildet, worin das sich schlecht einfügende Kind wohl unzureichend sein wird.

Man entdeckte, dass die Knaben bei allen untersuchten Altersgruppen den Mädchen in allgemeinen Kenntnissen überlegen waren: dieser Unterschied liegt zum Teil an den Unterschieden in der Umgebung ausserhalb der Schule und zum Teil an Beweggründen, die mit der Berufswahl zusammenhängen.

PSYCHOLOGY DOWN THE AGES.

By C. SPEARMAN. Two Volumes. (Macmillan and Co., 1937, pp. 288. 30s.)

So many reviews of this notable work have already appeared in various journals that it would be possible, and indeed profitable, to write a review of reviews. This is not the occasion for it, as it would be outside my commission. Yet there seems little point in repeating the laudatory comments of other reviewers.

Rather than catalogue the contents of forty-two intensely interesting chapters I will begin by giving the briefest possible outline of the plan of the work. The first volume is divided into three parts. Part A (What Psychology is About) consists of a masterly survey of academic philosophy. Already the informed reader can see the direction in which the author is speeding with a sure tread. "If psychology would perform its mission as an empirical science, it must at least be able to predict the future." In brilliant fashion it is shown how philosophical or physiological issues have continually arrested psychological discussion. To anticipate we may say that the final picture of psychology which Spearman has painted is that of a young and lusty, yet independent, science. This picture has been drawn by one who perhaps more than any other deserves the credit of emancipating psychology from being a bond-maid of philosophy. Yet withal he remains a true philosopher and accepts the sovereignty of philosophy in its own sphere in no niggardly fashion. He is concerned, however, that no philosopher should stop the psychological chariot by idle metaphysical speculations. "As regards the philosophical object, this, too, need not bother us, for the problems about it we hand over to the philosophers." Equally certain is he that the shortcomings of physiology cannot be regarded as a sufficient reason for calling psychologists to halt until the physiologists have set them on their feet. I will leave it to the reader to see how Spearman solves the dilemma. It must suffice now to say that Spearman is out to set psychology on its feet as an independent science without casting aside physiology as a broken reed or turning out philosophy with a pitchfork. He can thus claim to be regarded as a middle-of-the-road psychologist. In passing, I would note a tendency on the part of some authors to use this term "middle-of-the-road" in a slightly different sense, namely, to denote the psychologist who accepts well-established psychological facts irrespective of which school gave them birth and who then *organises* them into a science. This is all very well.

It all depends on the critical word *organises*. It surely must involve *interpretation* of those facts. The second part of the first volume is Part B (What the Psyche Can Do). It contains amongst other topics the most trenchant criticism extant of faculty psychology. The third and final part of the first volume is Part C (How the Psyche is Constituted). It is in many respects the portion which deserves the closest study and will probably give rise to many debates among academic psychologists. Spearman here switches his torch on many doctrines which concern the fundamental constitution of the psyche, including modern theories such as are associated with authorities like Wundt, Titchener, McDougall, Janet, Freud, the Gestaltists and many others. A reviewer might pick out a sentence here and there and attempt to score some trivial points, but a collection of such points would count nothing towards a "win on points." That sort of reviewing leads nowhere in particular. As one who long ago joined the ranks of the London School wholeheartedly the present reviewer cannot be expected to make adverse criticisms of any import. From that viewpoint his choice for the present task may be questioned. If asked to take the opposite side in an academic debate the most he could attempt would be to take up some less fundamental aspect of the noegenetic system, e.g., to query the number of different relations, or to ask if the scope of the relation of psychological objectivity is not too wide, or to suggest that it is inextricably intermixed with other relations, or that it is capable of being sub-divided into more than one class of relations. Criticisms on these lines would at least be constructive and would be welcomed by Spearman, who frankly acknowledges that his scheme of relations is not regarded as final and that "there is naturally room for difference of opinion." All the same, it is remarkable how little criticism has been made up to date.

The advanced student will be well advised to study Section C, as he will probably find it more difficult than most portions of the second volume. He will probably be acquainted already in a general way with the second volume in so far as he has mastered the well-known laws enunciated by Spearman, and has made himself familiar with the experimental conclusions of the London School as presented in Spearman's earlier works. But I hasten to add that the second volume is also entirely fresh in its arrangement and merits close attention if only for the fact that it brings the system up to date and shows how rich a harvest has been garnered in recent years.

The second volume is divided into two parts. Part D (What Follows What) begins with chapters on the Need of Laws and the Nature of Scientific Law, followed by chapters on each of Spearman's Laws: the

Law of Retentivity (Dispositions and Inertia), the Law of Control, the Law of Constant Output, the Law of Fatigue, the Laws of Noegenesis, the alleged Laws of Oresis, and the Laws of Basal Conditions.

Part E (What Goes with What) is concerned with the psychology of individual differences, with general, group, and specific factors, with the Two Factor theory, with typology, and with orectic factors. In this part it can be seen at once how vital and inherent a role is played by those mathematical concepts which were devised by Spearman, and which, no less than the noegenetic and other laws, will always be associated with his name. It would be difficult to over-estimate the importance of these discoveries for psychological progress. Spearman's system of psychology is one of several which were founded in this century, and it has steadily gathered force. Behaviourism is another system; it may be debatable whether it is a psychological system, yet there is no gain-saying the fact that it has enriched psychology in many ways. Yet apparently it has already passed its peak. That is very far from being the case in the system under review. In recent years its tide is sweeping into the orectic field and even in the study of character coefficients of correlation are continually being calculated. In addition, Spearman claims that many so-called orectic activities are really explained by his fundamental laws. Not only the sentiments but psycho-analytic mechanisms such as defence-reactions, compensations, sublimations, distortions, condensations, and displacements do not seem to require the invocation of any laws other than those enunciated by Professor Spearman.

Such then in the briefest compass is the plan of these two volumes. In boldness and originality of conception, in width and depth of scholarship, in charm and vigour of language they stand unique. Moreover, the clearness of the exposition and the wealth of illustrations enable even the general reader to gain some idea of the richness of the fare provided for him. I do not by any means suggest that *any* reader will not find many a stumbling block in the course of these eight hundred pages. That is bound to occur in a work so vast and ambitious, but it is abundantly clear that the author must have taken extraordinary care and patience to ensure that his discourse should be intelligible and he has succeeded admirably in his task.

In law it is well recognised that judges cannot function efficiently until they have practised law for the best part of their lives. In psychology also the same principle has its applications. Even the most implacable critic would have to admit that only a senior psychologist of the first standing could even attempt the writing of such a book as

Psychology Down the Ages. There is, however, this difference between law and psychology. There seems no lack of good judges, but, as I have suggested elsewhere, how many psychologists are there who by general acclamation possess the necessary qualifications for writing such a stupendous *magnum opus*?

It would be difficult to estimate the influence of this work on students of the mind. It will help them to assess the various systems which claim their attention. Thus Spearman directs some powerful criticisms against the Gestalt School, and students may grasp the significance of the points at issue. Again, these two volumes may help students either towards philosophy or towards psychology. I resist the temptation to perpetrate the dichotomy that students of the mind are either philosophers or psychologists and will be content to note that some like to study the nature of reality or of being while others prefer to investigate practical or ground-level problems. This work should aid them to discover their bent. Finally, the work will be welcomed most of all, perhaps, by those who have been members of Spearman's laboratory, who have been privileged to see the works from the inside as it were. Were these various laws educed or was it a case of intuition? There are passages in this work in favour of education. Thus Spearman refers to glimmerings in the writings of the Epicureans, Stoics and Empiricists, as well as to those of Stout, Hobhouse, Carveth Read, Meinong and K. Bühler. The criticism of Claud Bernard is also quoted with approval: "When one begins to base opinion upon feeling, upon inspiration, or upon a more or less vague intuition about things, one is outside of science." It will amply repay the student to look up the references to "intuition" in the admirable index which was prepared by F. C. Thomas. He will then realise that both system-builders and ordinary mortals agree as to the process. It is a question of more or less.

LL. WYNN JONES.

LA NAISSANCE DE L'INTELLIGENCE CHEZ L'INFANT.

By JEAN PIAGET. (Neuchâtel : Delachaux et Niestlé, 1935. 8 fr.)

THIS book marks a new phase of Professor Piaget's work. He now turns his attention from childhood to infancy, and in a series of three volumes, of which this is the first, he proposes to examine the origins in infancy of those features of childhood upon which he has already thrown so much light. The volume before us is concerned with the birth of intelligence. Volume II will deal with the manner in which the infant builds up his picture of the real world, and Volume III with the development of imitation.

The greater part of the book is taken up with a reasoned account of actual observations of three children ; this is preceded by a general discussion of the nature of intelligence from both the biological and psychological points of view, and followed by a concluding section in which Piaget compares various hypotheses of intelligence with that which he himself finds it necessary to adopt.

His thesis is this : intelligence is at once assimilation and accommodation. In his introduction he amplifies this, illustrates it in the body of the book by a mass of observations, and by way of conclusion shows its validity as compared with other hypotheses.

Intelligence is assimilation and accommodation. Piaget, by early training a biologist, begins by showing that intelligence can be described in biological terms ; that is to say, intelligent behaviour is only a special case of behaviour found in all living organisms. Fundamentally, we can describe all behaviour as having two features : organisation and adaptation. At any moment in the career of an organism we can consider either its structure or the manner in which it is adapting itself to its environment.

What is the nature of this adaptation ? asks Piaget. Both biology in dealing with the evolution of the species and psychology in dealing with the development of the individual attempt answers, presenting five parallel types of theory. Lamarckism and associationism both insist upon the rôle of environment without demanding any "internal activity" ; vitalism and "intellectualism" both attribute to the organism an innate faculty of dealing with its environment ; "preformism" and "apriorism" both emphasise the existence of structures prior to any experience ; mutationism and pragmatism alike maintain that adaptation

occurs by the environmental selection of characters appearing by chance ; while relativism in biology is paralleled in psychology by the view, here sponsored by Piaget, that adaptation is equilibrium between assimilation and accommodation.

Both of these terms are used in a broad way. By assimilation Piaget means the manner in which the organism tends to preserve its established modes of behaviour by subordinating its environment to them. Thus the mere taking of food is assimilation ; so too is the exercise of intelligence in that the organism attempts to bring new experiences within the framework of schemes of behaviour which it already possesses. By accommodation Piaget means the effect of environment upon the organism in causing it to depart from its existing modes of behaviour. Thus as soon as the organism begins to pursue its food it is accommodating itself to its environment ; and intelligent behaviour is accommodation in so far as the organism adopts new means in order to deal with unfamiliar situations.

Piaget then proceeds to illustrate his view of the equilibrium of accommodation and assimilation in intelligence by describing the development of his three children, a boy and two girls. Intelligence, he makes it clear, is not a mode of behaviour which appears suddenly at a given moment in infancy. What we begin to name intelligence, say towards the end of a child's first year, is in fact only a development and refinement of behaviour characteristic of the child from birth. This development takes place in two phases : sensori-motor adaptation gives place to intentional adaptation.

The former period, again, falls into two stages. Stage I—beginning at birth—is characterised by the exercise of inherited reflexes ; sucking, for instance. The child, endowed with relatively fixed modes of behaviour, proceeds to assimilate his environment to them. But even at this early stage, accommodation is not altogether lacking : the child adapts his sucking reflex to a variety of situations. And both assimilation and accommodation are determined by the child's needs.

In Stage II (which may begin at the end of the first month) simple acquired adaptations and "circular reactions" appear. The child, for instance, begins to suck his thumb ; and further, prolongs and repeats this sucking for the sake of the pleasure which it brings—"circular reaction," as J. M. Baldwin terms it. At this stage, once more, both assimilation and accommodation play their parts : the child assimilates the thumb as a stimulus to sucking and at the same time accommodates this reflex to the new stimulus. All this, again, is determined by the urgency of the child's needs.

With Stage III (from about the third month onwards) "secondary circular reactions" and intentional behaviour first appear. The child's activity begins to be engaged with the external world, and by circular reaction he prolongs the new experiences.

For instance, he wriggles about in his carriage and observes that toys suspended from the hood dance up and down. Here we have assimilation, in the sense that a new stimulus becomes the incentive to movement, and at the same time accommodation in that the child attempts to adapt his movements in order to secure the new experience. The advance on Stage II is that the stimulus gives rise to intentional movement.

Stage IV (from about the seventh month onwards) finds the child beginning to apply acquired modes of behaviour to new situations. He sees, for instance, a toy lying on the canopy of his cot. He wriggles, moves his arms, bounces about—all without result. Then he tugs at the cords of the canopy and when at last the toy falls, snatches at it. We may say that he is co-ordinating all these forms of behaviour, with the intention of securing the toy. At this stage assimilation consists in bringing a particular new activity (e.g. securing the toy) into relation with an existing scheme of acquired behaviour (e.g. tugging); while at the same time there is the accommodation of this co-ordinated scheme to the new situation.

Stage V (from about the eleventh month onwards) brings us to "circular reactions of the third grade," and the discovery by active experiment of new means to secure desired ends. Circular reaction takes the form of repeating the pleasant achievements of Stage IV; In this way the child will, for instance, learn that a stick or a string is a possible means of obtaining a desired object. Thus he assimilates the stimulus to the one valid and successful mode of behaviour while he accommodates the newly-discovered method to further situations.

With Stage VI (from about the end of the first year) the child at last reaches the point of inventing means to secure his ends, i.e., by mental activity he anticipates his physical behaviour. Piaget gives us an amusing example of this. He observed his daughter Lucienne at the age of sixteen months trying to get a watch-chain into the narrow aperture of a slightly opened matchbox. Having used various methods and failed, she sat for a while and contemplated the narrow opening. Then her mouth began to open, at first slightly, then more widely—she was miming the opening of the box. Then she inserted her finger into the slit and without effort opened the box.

With behaviour such as this we have reached what would generally be recognised as intelligent conduct ; but it is no sudden step. It has been led up to by progressive stages from the beginning of the child's life.

How is to be explained? By way of conclusion Piaget takes up each of the four hypotheses mentioned in his introduction and attempts to show their inadequacy in the face of the facts he has adduced. Associationism ; the hypothesis of fully-formed innate intelligence ; Gestaltism ; "trial and error"—each of these is inadequate in some measure. The only theory that will serve is that intelligence is essentially a balance between accommodation and assimilation.

The above outline can do no more than indicate the thoroughness and the richness of Piaget's treatment of his topic. In particular, he has put every student of infancy in his debt by the manner, at once profuse and cautious, in which he presents his observations. Many will welcome his support of a method which, though it goes back at least as far as Preyer, for a time seemed to be superseded by the statistical analysis of numerous cases, but which has in recent years again been productive of fruitful results—the serial observation of a few children over an extended period. Some, too, will welcome the evidence that, in infancy at any rate, intelligent behaviour does not always appear as a sudden new step.

At this point, however, one may venture on a criticism. No other method, perhaps, gives so much insight into the actual processes of development. But if general conclusions are to be drawn from observation of a few children, then surely it is of the first importance for the psychologist to compare his own results with those of others. There is to-day a steadily-growing body of observations of infant behaviour. To this Piaget has hardly referred ; still less has he attempted any systematic comparison. This, of course, in no way reduces the value of his observations. But since he compares his theory with other theories, one feels it would add cogency to his argument if he were also to compare his observations—even in a schematic fashion—with other observations.

It must further be added that at times Piaget does rather less than justice to the hypotheses which he rejects. Thus surely it is something of a *tour de force* to suggest that Lamarckism and associationism insist solely on the influence of environment and entirely ignore the internal activity of the organism.¹ Again, it is unlikely that Spearman would


¹ "La solution du *lamarckisme*, selon laquelle l'organisme est façonné du dehors par le milieu . . . à cette hypothèse . . . correspond en psychologie l'*associationnisme*, pour lequel la connaissance résulte elle aussi des habitudes acquises sans qu'aucune activité interne qui constituerait l'intelligence comme telle ne conditionne ces acquisitions." (p. 23.)

agree that his noëgenetic principles result in a picture of the organism *passive* in face of a reality which already possesses its own structure (p. 363). One would like to have seen Piaget attempting to assimilate other hypotheses to his own and accommodating his own to them.

One final remark is perhaps not out of place. No man has done more than Piaget himself to make clear to us the relation between language and thought in childhood. In the introduction to this volume he tells us that it is in the nature of an amplification of his earlier work on that subject. It is therefore all the more surprising to find throughout the whole of this book hardly a single reference to the beginnings of language. It may be that he will give some attention to this in Volume II. But even so, one misses the light that would have been thrown on the growth of intelligence by concurrent observations of the growth of speech.

All these remarks, perhaps, no more than express the fact that the reader, already given so much, greedily wishes for more. For there is no question but that its author has here presented us with a book for which every student of childhood will wish to offer him sincere thanks.

M. M. LEWIS.



OUTLINES OF RESEARCHES REPORTED IN THESES PRESENTED FOR HIGHER DEGREES OR DIPLOMAS.

THESE OUTLINES MUST BE SUBMITTED THROUGH THE HEAD OF
THE DEPARTMENT IN WHICH THE RESEARCH WAS CARRIED OUT.

A Study of the Threshold in relation to the Investigations on Subliminal Impressions and Allied Phenomena.

*Thesis approved for the degree of Ph.D in the University of London, from
University College, London.*

By R. P. B. K. PILLAI.

THE thesis contains a review of previous work on thresholds in all fields and brings this work into relation with the various theories of the "subliminal mind."

The experimental portion of the thesis is devoted to a study of subliminal impression by methods allied to those employed by Coover in *Experiments in Psychical Research*, 1917. The chief experiments were in the field of vision and hearing. In the visual experiments typewritten letters were exposed to the subject one at a time and preliminary thresholds for legibility were obtained in terms of distance between the subject's eye and the letter. The subject was then moved to a distance considerably beyond that which corresponded to his threshold. The letters could now be seen, if at all, only as the faintest dots or streaks. The subject proceeded to guess what letters were exposed. After twenty-five letters had been guessed (the letters being exposed in predetermined random order), the subject repeated the experiment four times, each time 6 inches nearer to the letters, though never near enough to allow of conscious discrimination of them.

In the auditory experiments the procedure was in general similar except that the letters were whispered by the experimenter with the teeth closed. In this series the subjects in the overwhelming majority of guesses were not aware of hearing any sound that could be taken as corresponding to the experimenter's whispers.

Forty subjects took part in the visual experiments, recording in all 10,125 guesses; sixty subjects in the auditory experiments, recording in all 13,500 guesses.

Statistical treatment of the results showed that in both experiments all subjects achieved scores significantly superior to those which they would obtain by chance. Individual differences are large and appear fairly consistent. There is little correlation between adeptness in the two experiments in the case of those subjects who took part in both. In both experiments there is a pretty steady increase in scores as the distance is reduced, so that it would appear that the subliminal influence is a function of the intensity of the stimulus. The sixteen subjects who repeated the experiments most often show no clear signs of improvement through practice. The correlation between supra-liminal perceptive ability and ability in subliminal guessing is barely significant, nor do the results indicate any significant difference as regards race, sex, age, or special training.

In both the visual and the auditory experiments some letters are guessed much more frequently than others, e.g., A, T, P, in the case of vision, T, S, P, in the case of hearing. In vision there is a low correlation between the number of times a letter is guessed and the proportion of these guesses that is correct—in hearing a higher correlation (.6). But even if we exclude the results for the most favoured letter (A, in vision, which gives 12.6 per cent of the total correct guesses for all letters, and T, in hearing, which accounts for a corresponding 11.4 per cent), the total number of correct guesses is still significantly above chance.

An Experimental Study of Humour.

Theses approved for the Ph.D. degree in the University of London, 1938, from University College, London.

By RABINDRANATH GHOSH.

METHODS.—Four different methods were used in the investigation: experimental; incidental observations; clinical observations; questionnaire study. The *Experimental Method* consisted in taking introspections and judgments (six-point scale from "Not Funny" to "Extremely Funny") from fifty-one subjects on fifty jokes (collected from various sources). The judgments were scattered more or less according to the normal distribution curve, but a fair amount of agreement amongst the subjects as to the "good" and "bad" jokes can be inferred. Sexual and national differences are not statistically significant, and seem to have been overrated by various authors. The play of inhibitions can be detected in the very rapid changes of judgments which occurred fairly frequently. The overcoming of inhibitions resulted in greatly increased appreciation. Attempts to justify and rationalise the inhibitions can be detected in the introspections. Analyses of the introspections permitted a classification of various factors, the most important being as follows: humiliation; aggression; retaliation; incongruity; ambiguity; violation of social or institutional authority; pre-existing feelings, negative and positive, to the subject-matter; expectation; self-reproach; anxiety; association; social sanction and censor; language difficulty. These factors have both negative and positive effects on the appreciation of humour. In some cases, factors have been overstressed or even read into the jokes, while other factors were neglected or overlooked in order that the jokes might be made to fit into the pattern of the subjects' mental states. Intrapsychic conflicts were sometimes externalised. *Incidental Observations:* 320 observations were collected from the everyday human behaviour. Four types of reactions were observed: mutual enjoyment culminating in laughter; retaliation by a counter joke; anger in the *joke-target* (at whom the joke is aimed); sorrow in being attacked. Sexual jokes need companies of equals. Apparently harmless jokes conceal their real aims and are very often disguised as riddles. In jokes against well-known men there is often a displacement of the joke-target. Fooling and leg-pulling, amongst many other forms of humour, have mutual seduction and indulgence as their motives. Amongst the many purposes which humour serves to satisfy are the following: escape from reality; expression of aggression; a mode of dealing with aggression; a mode of return to reality; recreation and entertainment; creation of friendliness; gratification of pride or re-establishment of self-regard.

Clinical Observations.—Material has been obtained from the psycho-analysis of a homosexual patient. Musical abilities combined with a sense and use of humour enabled the patient to work out on defensive lines a professional career with promise of sufficient emotional and social gratification. Humour afforded him a defence against guilt-feeling and consequent inferiority and at the same time provided a cover for his aggression.

Questionnaire Study.—Answers on twenty-five questions on various aspects of humour from twenty-three subjects were analysed and classified. The conscious reports abounded in rationalisations on the part of the subjects, yet a deeper understanding was possible on analysis. The underlying aggression in humour was perceived, as also the flight from reality. Humour was found very difficult to define. All the subjects with the exception of one experienced the appeal of humour, and appreciated its high social value. The exceptional subject experienced anxiety whenever she felt the occasion for humour and at the same time felt ashamed at her inability to appreciate it. Before producing a joke the subjects wanted to be reassured that it would not hurt any of the hearers. Tolerance of jokes was in some cases interpreted as a friendly gesture. Intimacy with an individual reduced the resistance to joke-making. The ability to make jokes and witty remarks lends charm to a personality. There was noticed a definite search for humour. The company of revered persons has a deterrent effect on the appreciation of humour. Definite attempts are often made to memorise jokes. Intimate friends are the ideal company in which to tell sexual jokes. A contradiction was observed between the subjects' social and ethical conceptions and their use of jokes.

CONCLUSION.—The common inferences from the four methods are the following: humour serves as an escape from reality, an expression of aggression, etc.; failure of humour is attended by unpleasant consequences, viz., the arousal of pity, anger, annoyance, sympathy; the search for humour is a very common phenomenon; humour affords a cover for the expression of many socially tabooed wishes; humour can make aggression, retaliation, and humiliation innocuous; sexual jokes are best enjoyed in the company of equals.

Moral Attitudes in relation to Upbringing, Personal Adjustment and Social Opinion.

Thesis approved for the degree of Ph.D. in the University of London, from University College, London.

By S. BRAHMACHARI.

FIVE questionnaires were given to 120 subjects between the ages of 19 and 35, most of them being senior students at the university, forty of them women. The questionnaires were: (1) G. B. Watson's on "Conditions of Upbringing" (slightly adapted), (2) Freyd's on Introversion-Extraversion (slightly adapted), (3) a new questionnaire of twenty-three items on Social Attitudes, (4) a new questionnaire of thirty-five items or propositions on Moral Attitudes, the subjects being also asked to estimate with regard to each proposition the "sources" of their attitude, how far they realised the proposition in their actual conduct, and how far any discrepancy between their ideal and their actual conduct caused them worry or conflict; (5) the Thurstone Neurotic Inventory. The first four questionnaires were answered individually in the

presence of the experimenter, the subjects being asked to record their answers on an eleven-point scale from +5 to -5 and being encouraged to make comments or supplementary statements. They were also asked to indicate the source of their social and moral attitudes, and to name a character from history or fiction which approached nearest to their ideal. The Neurotic Inventory was filled up at home in accordance with the three-point scale used by Thurstone.

Correlations between the various questionnaires as wholes revealed a positive correlation of .60 between strictness of moral attitude and actual conduct, but a negative correlation of -.61 between moral attitude and discrepancies between attitude and conduct, so that those with the higher ideals also acted (at least in their own estimation) more morally, though there was nevertheless a greater difference between ideal and actual conduct than in the case of those with lower ideals. The correlation between discrepancy and conflict caused thereby, though positive, was barely significant; and, as regards the individual items, many cases are reported of large discrepancies causing little conflict and small discrepancies causing much conflict. The marks and comments clearly reveal the presence in some cases of two levels of morality, a conscious and rational attitude which allows satisfaction unless it is clearly harmful, and a more unconscious and primitive attitude, which usually demands a stricter adherence to conventional codes. Conflict can arise between these two levels of morality, as well as between either of them and the demands of instinctual gratification. Experimental verification is thus provided for certain psycho-analytic findings with regard to the nature of the Super-Ego.

Moral factors and the demands of instinct are both operative in the choice of ideal figures from history or fiction.

Tendency to neurosis (as measured by the Inventory) shows a small negative correlation with actual moral conduct, a small positive correlation with discrepancy between ideal and conduct, and a moderately high positive correlation with the conflict caused by the discrepancy.

Tendency to neurosis also correlates highly (.63) with introversion; it appears indeed, that the Neurotic Inventory and the Introversion questionnaire have much in common. "Severe upbringing" shows moderate positive correlations with both introversion and tendency to neurosis.

"Personal observation" appears as the most important source of both social and moral attitudes, though, as secondary factors, friends and teachers play an important part in social attitudes and parents in moral attitudes.

The Social Attitude questionnaire was so arranged that it afforded what would appear to be a measure of general progressiveness or conservatism. This measure, as obtained from the whole questionnaire, showed no significant correlations with the other measures. Nevertheless, coefficients of association between the various items of the Social Attitude questionnaire showed that there was a very strong tendency for those who were more than averagely "progressive" or "conservative" as regard one item to behave similarly as regards other items. This confirms the findings of several American investigators, whose results point to the existence of a general tendency to radicalism or conservatism. So far as the present results go, there appears to be comparatively little difference as regards the saturation of the various items with the general factor of "progressiveness."

A similar treatment of the Moral Attitude questionnaire revealed a less strong tendency towards a general severity or laxity of moral standard. In view of the positive correlation with actual conduct, this tendency may be connected with Webb's "*w*" factor.

Motivation in Group Testing.

A Thesis presented for the degree of Ed.B., Glasgow University, 1938.

By P. D. BARRON.

MUCH of the statistical structure of Spearman's theory of intelligence is based on the assumption that intelligence is a cognitive function little affected by conation. The evidence supplied by Spearman does not seem to be very conclusive. A fairly comprehensive survey of former research on the subject suggests that, though much of it is of no particular value, yet, such as it is, it seems evidence against, rather than for, Spearman's view.

In this experiment it was decided to use motivation of strict discipline vs. no particular incentive. In order to control the effects the following scheme was applied: Four non-qualifying classes were taken from four schools in Glasgow. Each class had slightly more than forty members, the average age being 12 plus. Owing to absences, the average number of pupils to take both tests in each class was thirty-five. All the classes were somewhat backward, had reputations for comparative "toughness" and rowdiness, and were in charge of firm disciplinarians. The Cattell Test, Scale 2, was used.

CLASS I.—Form A was applied here with the class teacher and the experimenter present. The teachers' task was to ensure perfect order and work at the apparent maximum of the pupils' ability.

A week later Form B was applied to the same class, but this time the teacher was absent and the experimenter adopted the rather weak, apologetic attitude demanded by the experiment; whispering, listlessness, etc., being ignored.

CLASS II.—Form A was applied under the second set of conditions, and Form B a week later, under the first set.

CLASS III.—Forms A and B were given a week apart, both under conditions of discipline, with the class teacher and the experimenter present.

CLASS IV.—Forms A and B were applied a week apart, both under conditions of slackness, with the class teacher absent.

An attempt was made to have the two sets of conditions as similar as possible in the various classes. The scheme, as operated, was made intentionally extreme to exaggerate results (if any).

RESULTS.

	<i>Average Scores.</i>	<i>Sigma.</i>	<i>N (No. of Cases).</i>
1A (Discipline).....	32.37	14.51	35
1B (Slack).....	25.74	10.91	35
IIA (Slack).....	23.53	13.20	36
IIB (Discipline)	36.56	10.67	36
IIIA (Discipline)	36.42	7.89	35
IIIB (Discipline)	40.17	9.34	35
IVA (Slack).....	30.94	10.68	35
IVB (Slack).....	28.68	8.65	35

The average gain in score is, therefore, 9.8 approximately in Groups I and II under conditions of discipline, as compared with work under slack conditions. In the control groups III and IV the average difference is 3.

The superiority in scores under conditions of discipline proved to be significant.

The correlations between the total scores (i.e., r_{AB} in the four classes) were found to be:

I, .633; II, .512; III, .983; IV, .878.

Using Fisher's " z " formula, it was found that the difference between:

I and II (.633 and .512) was not significant.

II and IV (.512 and .878) was reasonably significant.

I and IV (.633 and .878) was reasonably significant.

I and III (.633 and .983) was decidedly significant.

II and III (.512 and .983) was decidedly significant.

IV and III (.878 and .983) was decidedly significant.

The high correlations in III and IV show that when conditions are the same in two applications of the test (whether the amount of motivation is large or small) the two applications, A and B, are measuring practically the same thing. But the capacities measured seem considerably different if the two applications of the test are made under different conditions of motivation, since in that case the correlations sink to an extent which is highly significant.

CONCLUSIONS.

(1) Poor control of the testing situation and lack of motivation does seem to reduce scores. It follows, then, that testing must be done under standardised conditions.

(2) Correlation between individual differences in scores is still found even under uncontrolled conditions. Under these slack conditions the tests are still measuring some kind of ability, but the ability is distinctly different from the " g " measured under controlled conditions.

(3) While it may be true that conation has less effect on intelligence tests than it has on educational abilities, some influence is shown.

(4) Whatever is measured under conditions of discipline is measured more reliably than what is measured under the other conditions.

An Inquiry into the Use of Children's Drawings as a Means of Psycho-Analysis.

A Thesis presented for the degree of Ed.B., Glasgow University, 1938.

By JANETTE R. McINTOSH.

In this thesis six children were discussed. They were three boys and three girls, ages varying from six to thirteen, and I.Q.'s from sixty-eight to one hundred and twenty-six (New Terman, Test L), who all showed maladjustment to some degree. They were given a form of analysis in which the drawings they made and related associations, and the interpretations of the drawings and associations played the major part. These six cases were selected from a larger group of seventeen children studied. Each child was taken by the analyst alone and wherever possible the parents were consulted too. One child came five times a week for a month, and the others less frequently. The analyst made friends with the child and then invited him to draw anything he liked. In most cases there was no difficulty in getting

the drawings, and the best patient did eighteen pictures. After the drawing the child was asked to tell the analyst about the picture and thus the associative material was obtained. The analyst interpolated the smallest possible number of questions in order to keep the child talking, questions which were as far as possible of a non-suggestive character. She also made interpretations where the drawings and associations seemed to justify them.

With most of the children the analysis was not carried on long enough to show very definite results, but all cases showed improvement and in several this was very marked.

This method would appear to have considerable possibilities of being useful in handling children and patients who are more inclined to draw and talk about their drawings than to carry out analysis by ordinary verbal free association.

The best patient in this group had a distinct interest in drawing and doing things with her hands, and advantage was taken of this in the treatment. It might be a specially suitable method for children in the latent period, and too old for analytic play techniques. It was remarkable how fully the children's problems were expressed in the drawings they made.

A Study of Zeros in Elementary Arithmetic

A Thesis presented for the degree of Ed.B. at Glasgow University, 1938.

By ISABELLA H. MACKAY.

AIM : To investigate by means of an error study :

- (1) Which method of teaching the use of zero achieved better results :
(a) as a placeholder, (b) drilled as a combination.
- (2) The relative difficulty of different kinds of combination.
- (3) The relationship, if any, between zero combinations by themselves and in graded examples.

SAMPLE.—There were 910 cases aged 8-9 (20 schools) from 11 counties in Scotland. Each child did the four tests, i.e., addition, subtraction, multiplication and division.

RESULTS.—(1) Multiplication of zeros is by far the most difficult operation.

(2) The *placeholder* method of teaching gave considerably better results than the *drill combination* method, but no account was taken of level of intelligence and the tests were given by different administrators. Further experimentation is necessary.

(3) The zero \times a numeral appeared to be much more difficult than numeral \times zero combinations, but a corresponding difference was not found in other operations than multiplication.

(4) The difficulty ranking of zeros as combinations was quite different from that in examples. The zero has its own difficulty and a "functional" difficulty according to complexity of the example.

(5) The difficulty of an example increases as the number of zeros increases.

(6) The positional difficulty of zero varies according to the operation, e.g., an intermediate position is most difficult in multiplication.

(7) Individuals vary in their ability to manipulate the zeros in the different operations.

Homework and Leisure Time Activities

Thesis presented as part of the qualification for the Degree of M.A. in Education (Birmingham University), 1938.

By NORMAN F. MILLINGTON.

THE Board of Education Pamphlet (No. 110, 1937) on Homework, recommended that the amount set in certain types of school should be reduced. If children are to have less homework how are they going to use their increased leisure-time? Should they be left to their own devices or should attempts be made to control this leisure? With these questions in mind a questionnaire was drawn up and submitted to 170 children in a mixed selective central school with a technical bias, ages from twelve to fifteen years. It aimed at ascertaining: Time spent on School Homework—Conditions under which it is done—Homework preferences and suggestions for improvement—How leisure time is used at present—Hobbies, societies, reading, cinema, etc.—Pupils' views on the Cinema, Wireless, Newspapers, Cheap Literature, etc., and the extent to which they are used.

The results can be summarised thus:

About two-thirds of the total children tested think homework is of value and should be retained, but the fraction decreases with increasing age; in the 12-13 year group 57 out of 76 favour retention; in the 14-15 year group, 13 out of 29 favour retention; 82 children think they are set the right amount of homework; 24 children think they are set too little.

The average time reported spent per week was five hours for the 12-13 year groups and four hours for the 13-14-15 year groups. These averages cover very wide extremes—"as little as possible" to eleven hours a week. Twenty-three children say that homework is of value because it can be done free from interruptions by teachers and other scholars, and give indications of a realisation that being left to your own initiative is really a valuable experience. Seventy children would like it abolished, sixteen of them are unable to concentrate because of "the wireless, the baby or the rest of the family."

In the great majority of cases the children's leisure time does not appear to be very profitably used, and, from their remarks, there seems to be a strong desire for this leisure to be controlled to some extent, but not by being given homework which is a continuation of schoolwork; they would like it to be a little different. The children prefer for homework tasks which enable them to use their creative powers, even if only to a very small degree: drawing graphs, making models and obtaining illustrations for science, history, and geography lessons; painting maps, listening critically to an item on "the wireless" for class discussion. Reading one item of news in several papers; gathering information from various sources in preparation for a debate or lecturette.

Seventy-seven out of 146 children gave "thrillers" and war films as their first preference for cinema programmes, and eighty-nine out of 123 prefer wireless play thrillers. The most popular item on the wireless is music hall variety—155 out of 163. For reading, first preferences were: thrillers and crime, 53; adventure, 61; both out of 154.

It is suggested that the "creative urge" could be used in a voluntary form of homework which would lead to more permanent life interests than the more academic type of homework.

BOOK REVIEWS.

Collected Studies of the Dionne Quintuplets: By W. E. BLATZ and others. (The University of Toronto Press. 18s.)

This is a collection of six papers on different aspects of the development of the Dionne Quintuplets, the result of a co-operative enquiry organised by Dr. W. E. Blatz and approved by Dr. A. R. Dafoe, their medical attendant, who writes a brief foreword. The book is full of interest for the student of infant psychology. The first paper is a biological study, by J. W. MacArthur and N. H. C. Ford, which gives ample reasons for regarding the quins as an "identical" set. Even such matters as finger prints and soles of the feet are given careful consideration, and "throughout the set the two hands of any member are less alike than is one of her hands like a hand of a sister. In sib comparisons the opposite is the case." This section is followed by numerous plates and diagrams, which also appear in other parts of the book. The second section—*Mental Growth*—is by W. E. Blatz and D. A. Millichamp. This paper has, as one gathers of all, been published before, as one of the University of Toronto Child Development Series, No. 12. Here the paging inconveniently begins anew, and so with each paper. As there are many tables and photographs with no headings, or only a very general one, reference is extremely awkward. The value as a permanent book of reference is also reduced by there being no index.

In considering mental growth we have to recall that the babies were at least two months premature. This, of course, affects all the early dates of achievement. The rate of acceleration in the early months shows that to some degree the children are "catching up." Much the greatest retardation is in the development of language, as commonly noted in twins. The quins were subjected to frequent applications of tests beginning chiefly with the Gesell tests, the Kuhlman and the Merrill-Palmer. After the nineteenth month the order of the five children in these tests remains constant, and the general slope of development is always the same. Furthermore, among the four phases tested—motor, language, adaptive and personal-social, the rank order was the same for all five children at the end of the tests, and with few exceptions the same at the beginning. This last statement has to be balanced by another which is difficult to reconcile with it, namely, that "the rank order of the values for the four phases changes about throughout the test period." Another difference was that "the periods of acceleration were not identical in time for the five children" (p. 12).

The third section—*Social Development*—is by W. E. Blatz, D. A. Millichamp and M. W. Charles. Here come the most interesting individual differences between the children. The study of social development was based chiefly upon ten minutes' observation at monthly intervals for the period from one year and nine months to three years. This, of course, was a very small amount for each month. Furthermore, the point on which notes were made were very formal, namely, the number of "initiative contacts towards another" and the number of "responses to initiated contacts."

Hence a friendly approach seems to be grouped under the general heading of social aggression. Generally, these children were found to imitate each other much more frequently than was the case with a control group similarly observed. The authors call attention to the fact that the "most unequivocal conclusion to be drawn from these data, without stressing too much their reliability, is the appearance of differences already so apparent between the five children, where one would, perhaps, have expected closer agreement." Elsewhere, however, the authors suggest that individual personality traits are largely due to social environment: for example, "It is Annette's high frequency of initiated contacts that undoubtedly attracts to herself the watching attention of the other four, which, in turn, must stimulate Annette to further aggressive social behaviour in order to satisfy the developed need for an audience." Yes; but all this was surely first due to Annette's frequent initiation of contacts; so that it is ultimately attributable to an innate difference.

The fourth section on "The Development of Self Discipline" also emphasises individual differences in emotional traits and in the attitude towards authority. Here, also, I am at a loss to follow the author's conclusion that these differences are due largely to social environment, simply because, as they say, "their hereditary origins are identical." This latter assumes that the substantial physical identity of the quins is accompanied by identity even in the minutiae of mental traits and impulses, of which of course we have no proof. Indeed later, on page 40, the authors actually name as one of the environmental differences "the inter-action of the personalities of each individual one upon the other"—which presumes previous individual differences! Incidentally, it is strange to see the statement (page 26) that emotion is regarded as being manifested "when, in the opinion of the observer, the situation is one with which the child is unfamiliar."

The last section deals with the development of language. The quins were already noted at the age of 1; 0 to be slow in vocal language development. This is usually the case among twins, and the authors regard it as an accentuated case. The gestures of the quins, on the other hand, were remarkably expressive. The quins are now being taught English as well as French, so one may look for interesting material on bi-lingual development in the near future.

Taken together, the observations and records are a remarkable achievement through the ready seizing of a unique opportunity.

C.W.V.

Foundations of Educational Psychology: By PETER SANDIFORD
(Longmans, pp. 449. 12s. 6d.)

The scope of this book is suggested in the sub-title "Nature's Gifts to Man"; a second volume is promised which will deal with the "Psychology of Learning."

In the introduction the comparatively slow scientific development of psychology is attributed to the absence of early experimental work and to the difficulties of introspection as a scientific method. An outline of scientific methods follows and a brief but clear account of the various schools of psychology and of the contributions of each is given.

The next hundred pages deals with the vexed question of heredity and environment. After showing the educational importance of animal psychology the author goes on to give a most lucid exposition of the "Theory of the Gene" and of Mendelism. Variation in human beings is discussed on these lines and the probability curve and correlation coefficient explained. The bearing of the normal curve on education is discussed and a plea made for some sort of compromise between the uniformity of the older education which ignored individual variations and the over-individualising of some modern schools which tend to ignore the social group. Although more stress seems to be laid on the influence of heredity, the environmentalist's point of view is not ignored, and the importance of the teacher's contribution in supplying a social heritage is stressed.

Chapter II deals with Individual Differences. The author considers geographical position, race, ancestry near and remote, sex, and maturation as possible causes of these differences and the results of tests and carefully controlled experiments are given. Differences in sensory capacity and special abilities are considered, as well as those in general intelligence, and reference is made to the educational provision for individual differences, including such systems as the Dalton Plan, the Winnetka Plan, and the Project Method.

The Foundations of Behaviour are considered next. A special section deals with the endocrine glands and another with the nervous system, some account being given of the work of Wever and Bray, Adrian, Sherrington and Lashley.

This leads, in the next chapter, to a discussion of "Non-Variable or Unlearned Behaviour: Reflexes, Instincts and Emotions." The chapter concludes by laying stress on the educational importance of the experimental work that has been and is being done on emotional expression.

The last two chapters deal with "Intelligence: its Nature and Measurement" and "Personality: its Nature and Measurement." In the chapter on personality the author gives an account of the physiological theory of personality of Berman and of the views of Freud, Jung and Adler.

The book contains a valuable collection of material, and a full bibliography is given at the end of each chapter. Diagrams and tables are included in each section and help to make the more technical parts clear to the student. The author's style is clear and easy to read, and the ground is covered with a thoroughness which provides a very firm foundation for further study. Controversial issues are set forth in an unbiassed and direct way, and problems which require further research are indicated. Professor Sandiford has given the advanced student of educational psychology a valuable text-book and reference book. M.C.P.

The Psychology and Teaching of Arithmetic: By H. G. WHEAT. (G. G. Harrop and Co., pp. x+591. 8s. 6d. nett.)

It is not enough for the teacher of mathematics merely to know the processes, the conversions or the substitutions which he handles so deftly. Unfortunately, as a practised mathematician, he is apt to forget that these have evolved from experience and that there is a psychology involved. This is closely wrapped up with the history of the evolution of number, and intelligent treatment of his topics by the teacher implies that he should recognise this and know what the terms he uses are based upon.

This book shows, first, the psychology of number through its history; it then goes on to methods of teaching, to discuss processes and convey a large amount of information for the teacher on the actual procedure of lessons. One finds how the retarded pupil can be helped and how some of the difficulties well known to teachers of arithmetic can be handled. The outline of a course is useful, too, as a general guide as to what may be expected of children. The American appears to put the acquisition of fractions slightly later than we do in England, possibly he is right; but it still is a matter for inquiry, though it is true that some of our examiners for secondary school entrance have already at times shown evidence of some heart-searching in regard to the age at which knowledge of fractions and ability to work them should be attained.

Having put a proper emphasis on the rôle of perception and the growth of general ideas in teaching arithmetic, it is surprising that the author has not indicated means for calculating volume as well as area. One cannot argue that volume is beyond the powers of the elementary school child, especially with an extended school age within sight. One realises he may not be in a position to evolve a formula in instances beyond regular rectilinear solids, but he can be put in the position to accept as quite rational certain formulae for volume and area which he can learn to use properly. Inductive teaching is within the scope of a work on psychological bases.

A lot of spade work has been done in the United States on the teaching of number and arithmetical processes. It is not surprising, therefore, that this book by an American, written primarily for American teachers, should contain a bibliography largely composed of reference to American publications. The problems shown as illustrations in the text naturally are set in American terms, but this need not trouble the English reader who will know how to interpret them, and derive benefit from the book as a whole.

Quintilian on Education: By WILLIAM M. SMAIL. (Clarendon Press, pp. xlviii+144. 4s. 6d.)

"For what they are, would that they were ours," Bacon is reported to have said of the Jesuits. After all, he was thinking of contemporaries. Is it not an impertinence, though, to speak of those who, living in the past, had ideas harmonising with the best in our days to say "they lived in advance of their time"? Did Quintilian, born nineteen hundred years ago, live in advance of his day? Had he, or Comenius, or Wyse been saved up for the twentieth century would there have been a structure of educational thought such as we know? It is very doubtful. Each educator has builded in his own way, in his own place, and time. Quintilian made his contribution, which as a published work must have influenced thought in the first century (A.D.), even if it was mostly lost subsequently until the fifteenth, then to be received as startlingly fresh and acceptable.

References are now made to Quintilian in most patronising terms ; it is regarded as wonderful that he could be an educational psychologist ; that he could recognise the importance of play, not alone for teaching methods, but as relaxation for youthful minds that have, perhaps, not the staying powers of the mature adult, and so on. Mr. Smail shows Quintilian to us as a practical man, claiming, it is true, a good deal for formal education, as his writings indicate in the first book of the *Institutio Oratorium*, but also holding that education in effect is what is left after we have forgotten everything we have learned. Culture, for this is what it is, is to be based on broad foundations ; there is a place for music and geometry for example in the education of the orator, but, above all, the orator must be a good man. Morality should be a first charge in the early training of him who is to become an orator.

These remarks touch only the fringe of what Mr. Smail has presented in this book ; there is an introduction which gives an idea of Quintilian's life and influence in his own and in much later days, with a slight picture of the period in which he lived. Then there are selections from Books 1, 2 and 12 of the *Institutio*, giving us in a translation of his own words the ideas of Quintilian himself. These show him as the practical man with no patience with those who had got so far from facts and things that they could not make their meanings, even if they had any, clear to others.

Quintilian had naturally much to say about the specific training of the orator which is not given in these extracts, but what has been selected is of general application and should attract the attention of a number of readers, teachers, and especially psychologists and education historians. A.P.B.

Social Interest : a Challenge to Mankind : By Alfred Adler. (Faber and Faber, 10s. 6d. nett.)

Man's place in society and how it may be achieved is the topic with which Dr. Adler has been concerned. If we can accept his premises we should find little difficulty in adopting his conclusions, but it is not always easy to admit proposals about children's early resentments which the author claims to have such far-reaching effects in later life. At the same time, the opinions that are put forward should not be dismissed lightly even if the grounds are not as yet firm for general conclusions. Some of them are highly probable ; one cannot fail to be interested in the rôles played by the inferiority and superiority "complexes" in man's successes or failures, in his struggle for supremacy in the world of things and in society. Inferiority is responsible for practically all success and superiority leads to failure. "It is always the want of social feeling," writes Adler, "... which causes an insufficient preparation for the problems of life." In no better way can this be exemplified than by the numerous references made in this book to early treatment of their children by parents, by the neuroses set up by maladjustments in the home, and by behaviour of pampered children, second children in the family, etc.

Adler emphasises the use to be made of individual psychology ; he looks for much help in the explanation of life as a whole, from clear understanding of the earliest memories of childhood ; so he leads one on to read of types of failures, of perversions, and the reactions of those in later life who have been pampered as children.

The last chapter of the book is devoted to a philosophy of life which lays stress on the moral side, but the person must seek his greatest approach to perfection in cosmic integration. One can never be perfect, but one can get nearer and nearer to perfection. Logically, the chapter is in its right place, but had it been first in the book instead of the last it would, in our opinion, have given better orientation to the writer's thesis.

Great credit is due to the translators for the way in which they have carried out their part of the work. A.P.B.

La construction du réel chez l'enfant : By JEAN PIAGET. (Delachaux and Niestle, Neuchatel and Paris, pp. 390.)

Prof. Piaget's excellent book is much more interesting to the philosopher than to the educationist. It is an account of how the young child constructs his world. Important as this process is, it is not one in which definite, purposeful education

plays any part. The young child does it all by handling his toys and watching the persons and things around him.

The main aspects of the world which we all construct in early childhood are four in number: objects, space, causality and time; and the book is a careful account of how each of these aspects is elaborated by the young child. Prof. Piaget argues that in each aspect the child begins from an ego-centric position in which he regards the world and all its contents as depending upon himself and his own will, and advances gradually to a position in which he recognizes that the world has an objective existence and laws of its own, and that he himself is only one element in it. It is, of course, impossible to get directly into contact with the mind of the young child; but his behaviour can be watched and from what he does we can infer what he is and what he is not thinking of. Prof. Piaget watched three children from their earliest days to the end of their second year; and his conclusions are all supported by a long series of very careful observations. Those who are interested in modern controversies about the nature and growth of human knowledge will find much worthy of their study in this book.

So You're Going to a Psychiatrist: By ELIZABETH I. ADAMSON. (Pitman, pp. 263. 12s. 6d.)

Here is yet another of the many books which are being written by medical men and women, but which deal largely with psychological problems.

The physiological and medical elements in the book are no doubt sound (the reviewer is not able to judge them, though familiar doctrines are clearly in evidence), but questionable psychological statements are given in dogmatic style. Such statements as "most children are," etc., appear without any evidence or references to evidence except to a few particular cases. Presumably it is expected that the doctrines will be accepted chiefly on the ground of the medical qualifications of the writer, but it hardly seems adequate.

The Maturing Mind: By T. H. PEAR. (Thomas Nelson and Sons, Ltd., pp. 152. 2s.)

This little book, the first of a series of "Discussion Books," is evidence that psychology can be made as interesting as a novel. It is written primarily to be used by discussion groups, and deals with such topics as—the wisdom of being educated; acquiring background; the conversation and its delights; broadcasting and adult education and the mature personality. It is particularly encouraging to the middle-aged, for Professor Pear quotes psychological experiments upon adult learners which show that "adults can acquire many important subjects easily and rapidly and could learn much more than they do." Professor Pears' witty style is always a delight.

Vocational Training for the Deaf: By JOHN SPALDING, Head Master, Royal Residential Schools for the Deaf, Manchester. (Published as an appendix to *The Handicap of Deafness*, reviewed in the last issue).

(NOTE.—The term "deaf" is used in these notes in a restricted and special sense. It signifies those who, as children, were so deaf as to require education in a special school for the deaf.)

As the author reminds us at the outset, deafness is a *continuing* handicap. Education, which ceases so far as "school" is concerned at the age of sixteen, is a palliative only—not a cure—and the deaf child is still deaf when he has to face the world. This may appear to be an emphasis of the obvious, but experience shows it to be not unnecessary. Too many are ready to assume that at sixteen a miracle should have been completed, and the deaf child be entirely self-sufficient. This is far from being the case. However well he may have surmounted his handicap, he will always have difficulty, often in a considerable degree, in communicating readily with those around him. When he comes to seek a livelihood and wishes to

learn a trade, this difficulty looms very large. As Mr. Spalding is careful to stress, a useful occupation is essential to a life of satisfaction—as much a necessity to the deaf as to anyone else: and his education must make special preparation for that, simply because, being deaf, he meets with difficulties that can only be overcome by special methods, particularly of instruction.

Although there are brilliant exceptions it is inescapable that the greater part of the deaf must seek a living by the work of their hands. On this basis the author compares the training of a hearing child and a deaf child. The elementary education of the former goes on to the age of fourteen (perhaps to fifteen now, in some cases) without any suggestion that he should have specific trade training during his elementary school life. Why should the deaf child, whose schooling ends at sixteen and is a vastly harder task, have it curtailed for the sake of trade training? This has been done in the past and indeed, is actually done to-day in some places. There is, of course, a good deal to be said for giving the deaf a practical start before school life ends: but after careful examination of all sides of the question the author comes down solidly for no specific trade training during school life, that is, before sixteen, a position endorsed, I think, by the vast majority of teachers of the deaf.

Vocational training after sixteen is a very different matter. It seems to be generally agreed that if a deaf child can secure a place in industry at sixteen, where he will be properly taught a skilled trade, that is a perfectly satisfactory course; but taking a blind-alley job on leaving school (usually for the immediate financial relief of the parents), will just lead to unemployment later, and produce a morose and miserable individual. The solution advocated is that every deaf child who cannot at once be well placed in industry on leaving school, should have the opportunity, *of right*, of a proper training at a skilled trade. Such training is to-day available at various schools, as a full-time continuation course under workshop conditions; but there are insufficient vacancies for the whole country, and as the child's legal right to education ceases at sixteen, *his attendance at such a course is dependent on agreement between the local education authority and the parent*. It is contended that this training should not be the subject of caprice, that it should be available in every suitable case, of right. Much can be said in justification of the plea, but one point that will appeal generally is that students trained in this way have little difficulty in securing work, and in most cases are thoroughly established for life.

In the short compass of a dozen pages or so, the author has given an able and lucid exposition of the many problems involved in the vocational training of the deaf, and he brings cogent reasons in support of the solutions he advocates. From a wealth of discussion we take only one further point. At present no one, apart from the parent, has any legal responsibility for a deaf child after the age of sixteen. Apart from the voluntary work of the missions to the deaf, and voluntary help from his teachers too, there is none, as a rule, to give him a helping hand in difficulties, or to see that he is properly guided upon a course that will make his life a reasonable success. Mr. Spalding pleads that the responsibility of the State (which ultimately means local authorities), for the education and training and presumably also for the welfare of the deaf, should be extended beyond the termination of elementary schooling at sixteen, to cover the next four years or so, in order to ensure a really satisfactory start in life.

Mr. Spalding's excellent presentation of the pros and cons of vocational training for the deaf should be specially valuable to local education authorities in particular, in their efforts to alleviate the "handicap of deafness" for children.

F.W.C.

The Purposes of Examinations. Reprinted from *The Year Book of Education*. (Evans Brothers, pp. 146. 3s.)

In our educational system to-day examinations play a decisive part. The growth of free education, the rise of the modern universities, the increasing standardisation of the methods of entry into the professions, have all contributed to lift examinations to a position and importance such as they have never had before. But recently there has been apparent in many quarters a feeling of uneasiness lest

examinations may have become too important, or have failed to achieve the aims for which they were designed.

Numbers of books on this theme have appeared in the last five years, some of them seeking to assess the value of examinations by following up the careers of the examinees to see if their subsequent successes or failures bear out the verdict of the examiners, others examining the examinations themselves, and trying to ascertain what kind of qualification the student who has passed them may claim to possess.

The Purposes of Examinations, a symposium of contributions by many learned and distinguished authorities, is an example of the latter type of enquiry. It gathers together into one volume of modest size a statement of the relevant facts about a number of examinations which are taken by large groups of candidates at the present day. The book opens with an introductory survey by the general editor, Sir Philip Hartog: then follow a chapter each on the Free Place Examination and the Examinations for the Bar, three chapters on some of the best-known Degree Examinations of the University of Oxford, four on Degree Examinations of the University of Cambridge, and thirteen chapters on similar examinations in seven other Universities of England and Scotland.

Each chapter is written by an authority of special competence in his own field. That on the Free Place Examination is by Dr. Ballard, the Bar Examinations are dealt with by the Recorder of Cambridge, and the chapters on University Degree Examinations are by Professors or Directors of the departments concerned. In each case the general purpose of the examination is explained, with details of the conditions of admission to it, of its practical working, and the methods of appointing examiners and allotting marks. The contributors were not confined to bare statements of fact, but were free to give their opinions also, and most of them made use of this freedom.

It is these expressions of opinion, indeed, which give the book much of its interest and value; for, unlike the bare facts, they cannot be found in the ordinary works of reference. Two questions have been well to the fore in the minds of most of the writers. First, whether the examination under consideration is in itself a success or a failure, in relation to its avowed purpose. Second, whether the published results represent a fair verdict on the candidates. On the whole, the answers given to both questions are favourable, but there are some exceptions. Thus the writer on the London Internal B.A. General Degree confesses that in his opinion the examination has failed of its object. In another case the writer is uneasy at the large percentage of candidates for a medical degree who fall by the way and never reach the final stages of their examinations. He suggests that some alteration in the regulations may be needed.

On the question of the validity of the results as a verdict on the examinees, the writers on the University Degree Examinations are generally satisfied. But Dr. Ballard is frankly critical of the results hitherto obtained in the Free Place Examination. However, he believes that the examination is now moving towards a wiser régime.

It is, perhaps, a little surprising that a book with this title did not include chapters on the principal secondary school examinations. But those whose main interest is university education will find in it a wealth of information and authoritative comment; and it is no mean achievement to have gathered so much into a volume of such moderate compass.

Scholastic, Economic and Social Backgrounds of Unemployed Youth:
By WALTER F. DEARBORN and JOHN W. M. ROTHNEY. (Oxford University Press, 1938, pp. xi+172. 8s. 6d.)

The authors of this report are to be congratulated on the publication of an extremely interesting addition to the series of *Harvard Bulletins in Education*. The report describes an investigation into the background of young people to see whether any factors emerge which might differentiate unemployed from employed youth. The first section of the book is occupied with descriptions of the source of the data and the methods used in their examination, and the reader cannot fail to be impressed by the detailed care with which the investigation was carried out.

age is receiving attention and when the full account of what is done to meet the problems it raises is to hand one may rest assured that it will provide interesting reading. In the meantime, what has been done by Cadbury Bros. and some other firms remains as a stimulus to socially-minded employers of labour on a large scale elsewhere.

Statistical Tables for Biological, Agricultural and Medical Research: By R. A. FISHER and F. YATES. (Oliver and Boyd, London, 1938, pp. 90. 12s. 6d.)

The authors of this set of tables are to be congratulated on a laborious and well-executed piece of work which should be of service to all research workers in the biological sciences. These tables have been compiled with a more extended range of problems in view than those of the educational psychologist, and we have not so far found in our science any need for the Latin squares shown in Tables XV and XVI which have been of great value in agricultural research. On the other hand, the statistical problems of ranked data have especially arisen in educational psychology and Table XX for replacing ranks by normally distributed standard scores, and Table XXI for facilitating the calculation of sums of squares from such data, will supply a real need. Many educational psychologists have found the tables in Fisher's *Statistical Methods for Research Workers* exceptionally convenient for practical use. These are all included and somewhat extended. There are also factorials which are not generally to be found in the mathematical tables in common use, as well as the more usual logs., squares, etc. Table XXXIII is an extensive series of random numbers which are often required in experimental work. Educational psychologists should find this set of tables of real value in their work.

R.H.T.

Backwardness in School—its Diagnosis and Treatment: By P. A. BARONS. (Blackie, pp. vii+182, illustrations 7½-in.×5-in. 3s. 6d.)

This little book serves as a popular introduction to a vast and complex subject. The writer gives little that is new, but re-states the various aspects of the subject in an interesting way, and draws upon personal experiences to illustrate his points. Although there is little attempt to provide experimental evidence for the generalisations given, the treatment is clear, the language simple, and no knowledge of psychological terminology is assumed in the reader. After discussing in brief the causes of backwardness, and emphasising their complexity, Mr. Barons gives helpful suggestions based upon his own teaching experience to the teacher at the infant, junior and senior stages. These suggestions are not sufficiently detailed to be of complete assistance to the class teacher, but serve to indicate the lines upon which treatment should be based. Diagnostic and remedial work in the basic subjects is mentioned as necessary, but how to set about it is not fully discussed. References are given, however, and a long bibliography points the way to further reading. The teacher of a backward class, disheartened by the apparent futility of his efforts in the classroom, will find in this book encouragement to view the problem in a fresh light, and will certainly be cheered by the lucid and practical advice Mr. Barons gives.

The Feeding Behaviour of Infants: By A. GESELL AND F. L. ILG. (Lippincott, pp. 200+viii. 21s.)

There is little doubt that this is the most comprehensive and detailed study of infants' behaviour in the process of feeding that has yet been published. The book is, of course, of interest almost entirely to the advanced student of infant psychology, or to the medical man, though there are a number of important points which are very relevant to general genetic psychology, some of which the authors discuss in the concluding chapters on growth and on the adult-infant relationship.

The volume is freely illustrated with excellent reproductions of photographs showing almost every stage and aspect of infant behaviour while feeding.

The authors consider that excessive importance is attached to weaning by some psycho-analysts, though they recognise the danger of over-severe discipline in connection with feeding and toilet habits.

Their discussion of infant growth reveals that they look forward to important developments in biochemistry as a means of influencing infant development.

Characteristics of Good and Poor Spellers—a Diagnostic Study: By D. H. RUSSELL. (Teachers' College, Columbia University, pp. 103. \$1.60).

The problem of the teaching of spelling is one which needs urgent attention in our schools to-day, and any investigation which tends to help in its solution is very welcome. Such an investigation is reported in this book. It was carried out on sixty-nine individually-matched pairs of children selected from four New York public schools. Good and poor spellers were compared on eighteen diagnostic tests involving constitutional, academic-achievement and study-method factors related to spelling ability. Although the number of children tested was small, care is taken in presenting the results to show which are reliable and which are not so. The diagnostic tests used are given in the book and should prove of great interest to any teacher who is concerned with the teaching of spelling. Undoubtedly it is by the study of the results of such investigations as this that satisfactory diagnostic and remedial methods in the teaching of spelling will eventually emerge for the use of the class teacher.

A General Selection from the Works of Sigmund Freud: By JOHN RICKMAN. (Hogarth Press, pp. 329. 5s.)

This is a useful selection, the papers covering the period from 1910 to 1926. Dr. Rickman explains that no attempt has been made to include either papers giving case-histories or other evidence of the theories, or on the technique of therapy. With these necessary limitations the book provides a well-balanced selection of characteristic expositions.

A word of special praise is due to the editor for the excellent index and glossary. This is very full and very clear. It contributes almost a psychological dictionary of Freudian terminology, and should be most useful to a serious student of Freud's works. C.W.V.

Seven Stages of Childhood: By ELIZABETH SLOAN CHESSEY. (Herbert Jenkins, pp. 320. 7s. 6d.)

This book is written confessedly chiefly for the young mother who wishes to know something about the bringing up of children, and it is for such that it is likely to be of use. Such psychology as appears from time to time is of a popular and elementary type, but, on the whole, not unsound. There is much useful advice on physiological matters. There is, however, little of value for the serious student of psychology. Nor will it make much clearer to the harassed parent whether most mental disturbances, if they have to be dealt with, are really psychological or physiological in origin.

Adult Attitudes to Children's Misdemeanours: By HELEN BOTT. (University of Toronto Press, pp. 21. 1937.)

This records a simple but interesting enquiry as to the attitudes of five groups of people (teachers, parents, public health nurses, social health workers and mental hygienists) to various kinds of childish misdemeanours. The most interesting results were the following: teachers ranked destructiveness, disobedience, uncleanness, profanity and impertinence as important; mental hygienists regarded these as relatively unimportant. Mental hygienists regarded avoidance of the group, day-dreaming, shyness, lack of application, enuresis and mannerisms as important, whereas teachers regarded these as distinctly less important.

Personality Development in Children: By E. J. CHAVE. (University Chicago Press, pp. 348. 11s. 6d.)

This is a discussion of the training of children, bearing in mind the findings of biology, psychology and sociology, and with special reference to religious education. The author is the Associate Professor of Religious Education in the University of Chicago.

Guiding the Child: By ALFRED ADLER. (George Allen and Unwin, pp. 268. 7s. 6d.)

This is a cheaper edition of a book first published in 1930. For the benefit of readers unfamiliar with the first edition we may say that it is a collection of papers by various workers connected with the child guidance clinics in Vienna, Berlin and Munich, which were conducted according to the tenets of Adler's individual psychology.

Adler himself edited the volume and wrote several of the papers. Some of these deal with special cases, others are apt to be somewhat scrappy and vague. It is a great pity that in this new edition no index has been added.

OTHER PUBLICATIONS RECEIVED.

Oxford Historical Note-Book: R. A. Mears, pp. 32, 1s. (Oxford University Press.)

French Vocabulary and Syntax: G. H. Rochat, pp. 107, 2s. (Macmillan and Co.)

Saint Matthew: B. K. Rattey, pp. 192, 2s. 6d. (Clarendon Press.)

Introduction to Drama: G. Newbold Whitfield, pp. 186, 3s. 6d. (Oxford University Press.)

A Middle School English Course: F. T. Wood, pp. 184, 2s. 6d. (Macmillan and Co.)

The Merchant of Venice: edited by R. W. Fletcher, pp. 192, 2s. (Oxford University Press.)

Latin (for School Certificate): R. E. and A. E. Burns, pp. 141, 2s. 6d. (Macmillan and Co.)

Meaning and Style: A. F. Scott, pp. 190, 2s. 6d. (Macmillan and Co.)

Modern Essays in Criticism: edited by A. Caincross, pp. 301, 2s. 6d. (Macmillan and Co.)

Teaching of Religion: E. F. Braley, pp. 154, 3s. 6d. (Longmans, Green.)

Short Historical Plays: edited by E. R. Wood, pp. 237, 2s. 6d. (Macmillan and Co.)

The Aeneid of Virgil (Dryden's Translation): selected by B. Pattison, pp. 249, 2s. 6d. (Macmillan and Co.)

The Economy of Britain: a History: Croome and Hammond, p. 321, 6s. (Christophers).

THE SCIENTIFIC INTERESTS OF SENIOR
SCHOOL CHILDREN.

BY R. RALLISON.

- I.—*Introduction.*
- II.—*Method of investigation.*
- III.—*Are children interested in science?*
- IV.—*In which sciences are children interested?*
- V.—*Science topics most interesting to seniors.*
- VI.—*Influences on the scientific interests of seniors.*
- VII.—*Summary of conclusions.*
- VIII.—*References.*

I.—INTRODUCTION.

RECORDS of the investigation of the interests in science of children at the infant and junior stages already exist. Two parents¹, in the *Scientific Interests of a Boy in Pre-school Years*, show that genuine scientific interest first manifests itself at about four years of age². H. R. V. Ball³ has supplied results from junior children. C. Washburne, the American psychologist, has worked out the scientific interests of American children and based on his research a textbook⁴ for science teaching in secondary schools. This investigation concerns English senior school children between the ages of eleven and fourteen; it is an attempt to reveal the nature of their scientific interests.

Evidence has been collected and summarised to illuminate six aspects of the problem of scientific interest, viz.:

- (1) What is the relative distribution of interests between the scientific and non-scientific fields?
- (2) What is the "pattern" of the interests shown by children in the various science branches?
- (3) Which topics seem most interesting to children?
- (4) What are the typical questions that characterise the ages?
- (5) What is the influence of age on scientific interests?
- (6) What is the influence of environment on scientific interests?

II.—METHOD OF INVESTIGATION.

The following instructions were cyclostyled and, with a covering note, sent by post to forty school departments in the City of Newcastle-upon-Tyne and to ninety-seven schools in the remote rural areas of the County of Northumberland; between Newcastle and the rural parts of

The questions on non-scientific topics were simply counted; a supplementary investigation is to analyse fully their content and indicate their nature. On the numbers of boys and girls mentioned above, it was found that 4,931 questions from boys and 12,333 questions from girls came in this category.

III.—ARE CHILDREN INTERESTED IN SCIENCE?

After investigations had proceeded for some time with city children, it was realised that by counting the questions and references to non-scientific topics, instead of merely rejecting them, some information as to the extent of the interest of children in the science field, compared with their interest in other directions, might be obtained. All the interest sheets for township and rural children were counted in this dual way for scientific versus non-scientific influences. Table II summarises the results obtained; it gives the relative numbers of questions collected in the non-scientific and scientific fields; from these figures were calculated ratios "non-scientific/scientific" to indicate the distribution of the interest-field for each type of children.

TABLE II.—"NON-SCIENTIFIC/SCIENTIFIC" DISTRIBUTION OF INTEREST.

Age.	CITY.					
	Boys.			Girls.		
	N.Sc.	Sc.	Ratio N.Sc./Sc.	N.Sc.	Sc.	Ratio N.Sc./Sc.
11+	203	1824	0.11	1698	1095	1.6
12+	122	1248	0.10	959	696	1.4
13+	107	980	0.11	1285	924	1.4
	TOWNSHIP.					
	N.Sc.	Sc.	Ratio N.Sc./Sc.	N.Sc.	Sc.	Ratio N.Sc./Sc.
	N.Sc.	Sc.	Ratio N.Sc./Sc.	N.Sc.	Sc.	Ratio N.Sc./Sc.
11+	350	670	0.52	1116	641	1.7
12+	1078	1749	0.61	1890	1322	1.4
13+	636	1511	0.42	1671	1486	1.1
	RURAL.					
	N.Sc.	Sc.	Ratio N.Sc./Sc.	N.Sc.	Sc.	Ratio N.Sc./Sc.
	N.Sc.	Sc.	Ratio N.Sc./Sc.	N.Sc.	Sc.	Ratio N.Sc./Sc.
11+	666	1060	0.63	957	537	1.8
12+	1101	1332	0.73	1347	546	2.4
13+	666	1239	0.54	1410	846	1.6

Note: N.Sc., Non-scientific; Sc., Scientific.

Inferences made from the study of these figures are :

- (1) A great difference exists between the sexes with regard to the distribution of interests ; boys seem to concentrate their interests in the scientific field ; girls are much less interested in scientific topics—their interests are much more widespread around a definite, small core of scientific interests.
- (2) Environment seems to have a much greater influence on the distribution of boys' interests between the non-scientific and scientific fields than is evident with girls ; city boys have the greatest proportion of interests in science ; township boys come second to them in this respect ; rural boys have, for boys, the least proportion of interests lying in the field of science, although their interests are nothing like as widespread as was found for girls. Girls from all types of environment are remarkably similar in the constancy of their interests, tending mainly towards the non-scientific. Evidently, environment has a much more potent effect on boys' interests than girls'.

The first inference is supported by Shakespeare's results in *An Enquiry into the Relative Popularity of School Subjects in Elementary Schools*⁵, where is pointed out, ". . . a scientific outlook seems evident in the boys' lists . . . girls seem to maintain or develop an interest in reading, composition and literature . . ."

Similar confirmation is obtained from Hoy's "*An Enquiry as to the Interests and Motives for Study among Adult Evening Students*."⁶ The question ". . . would you like to attend a class in science treated in a popular way as if affects your everyday life ?" was put to find out the opinion in regard to this " rather undeveloped branch of adult education." "Seventy-eight per cent of the replies were in the affirmative . . . the demand for science among women was very pronounced."

The ratios, "non-scientific/scientific," proved of value as they afforded a bridge over which comparison became possible with Ball's⁷ results for juniors. From his data⁷, it was possible to calculate a value for the ratio, "non-scientific/scientific." A value of 0.5 resulted. As this value is of the same order (in the mathematical sense) as the values given in Table II, it seems that the interests at the senior and junior stages are similar as regards their distribution into non-science and science.

IV.—IN WHICH SCIENCES ARE CHILDREN INTERESTED ?

The number of times children in an age-group asked questions which on analysis could be included in a science-branch was taken as an index of the total interest of the age-group in that science-branch.

Thus, the 11+ group of city boys asked 859 questions with a biological significance out of a total of 4,121 questions. From similar sets of raw scores for questions from each type of environment dealing with electricity, chemistry, light and the other science-branches, it was possible to calculate the percentage distribution of scientific interests with reference to the different science-branches. Table III summarises the results obtained.

TABLE III.—PERCENTAGE DISTRIBUTION OF SCIENTIFIC INTERESTS.

Science Branch.	City.			Township.			Rural.		
	11+	12+	13+	11+	12+	13+	11+	12+	13+
Boys :									
Biology	21	23	18	31	29	28	38	36	38
Electricity	17	14	18	15	8	20	5	7	11
Chemistry	17	21	18	12	13	16	11	12	14
Mechanics	13	12	15	10	14	19	16	15	14
Light	6	6	8	3	2	1	2	4	3
Astronomy	6	6	6	6	6	3	6	4	4
Heat	5	5	5	4	8	1	5	8	3
General (Topics difficult to classify)	5	5	5	14	17	7	13	11	6
Geology	4	2	2	3	3	5	2	2	4
Sound	4	3	3	2	2	1	3	2	4
Magnetism	1	1	2	0	0	0	1	0	0
GIRLS :									
Biology	31	24	35	24	32	37	39	41	44
Electricity	10	14	15	11	8	10	5	5	6
Chemistry	26	25	18	14	15	21	23	19	19
Mechanics	9	8	5	19	13	7	7	9	6
Light	5	6	5	5	3	4	2	3	3
Astronomy	4	6	8	7	5	7	2	3	6
Heat	4	5	3	6	3	7	3.5	1	5
General	3	6	5	2	7	3	6	13	3
Geology	4	1.5	4	9	11	4	0.6	0.5	2
Sound	4	5	2	3	3	1	13	5	5
Magnetism	1	0.9	0.6	0.6	0.3	0.1	0	0	0

It is possible to make some comparison between these results and results from other relevant investigations. From Table III, *Scientific Interests of a Boy in Pre-school Years*⁸, have been calculated the following percentage distributions for spontaneous scientific interests over the age-range, two to five years :

Biology (sum the given data for animal life, self, babies, physiology, and plant life)	33%
Chemistry	10%
Mechanics	7%
Astronomy	15%
Physics (compare with the aggregate of results for electricity, light, heat, sound, and magnetism in Table III above) ..	27%

The general spread of the "pre-school" results is quite similar to that obtained for senior children in all districts. Recalculations were also made from Ball's data⁹ for the junior stage. This source gives about 32 per cent of junior interests in the biological field, which is quite comparable with the biological results contained above in Table III.

A suggestion that seems to arise from these comparisons of the distribution of interests in sciences at the pre-school, junior, and senior stages, is the predominance throughout of an interest in biology. For American children, however, Washburne¹⁰ found that biology came fourth in order of "interest"; physics came first, astronomy second, and physical geography third.

Study of Table III itself reveals certain tendencies; these are:

- (1) Four science-branches dominate the interests of both sexes, viz.: biology, chemistry, mechanics, and electricity.
- (2) Boys and girls are interested in the same sciences.
- (3) The character of the kind of science they are interested in seems to depend to some extent on the environment, particularly with rural children.
- (4) City and township children seem to possess a rather more widespread interest in a larger number of sciences than rural children, but the rural children seem to compensate for this by a more concentrated interest in a fewer science-branches.

V.—SCIENCE TOPICS MOST INTERESTING TO SENIOR CHILDREN.

It has been proved that "Learning is much more effective if preceded by questions which let the child's mind forecast what he is going to learn."¹¹ To enable the science teacher to apply this by using the questions about scientific topics that naturally occur to children, a selection is now to be given of the more typical questions. The soundness of starting from the child's own questions is pointed out by three recent influential sources, viz.: Board of Education *Handbook of Suggestion* says, ". . . the questions that arise spontaneously in their minds . . . should be the starting point for the teaching"¹²; again the same phrase, this time applied to the middle forms of secondary schools, appears in the Board of Education Pamphlet 114¹³; finally, the Science Masters' Association advocates in *General Science* an attempt being made to "try to teach a child the things he wants to know."¹⁴ As a preliminary to being able to do this, they recommend finding the questions a boy wants answering.

CLASS I.—UNIVERSAL INTERESTS.

The following questions were found to be outstanding in arousing the curiosity of children of all ages (boys and girls) in the three kinds of environment used :

BIOLOGY INTERESTS.—*Human body*. How hair grows? ; how we breathe? ; how eyes work?

Animals. How do fish breathe? ; how do fish swim? ; about worms ; how hen makes eggs? ; about insects ; how do birds fly? ; about birds' habits ; how is chalk made? about wild animals.

Plants. Why are flowers scented? ; why are flowers coloured? ; how do trees grow? ; how do plants grow?

ELECTRICITY INTERESTS.—How does wireless work? ; how does telephone work? ; cause of thunder and lightning ; about electricity.

CHEMISTRY INTERESTS.—How is paper made? ; why is the sea salty? ; how does a fire burn? ; how is ink made? ; how is glass made? ; how is coal gas made?

MECHANICS INTERESTS.—How does an airplane fly? ; how do clocks work? ; how does a bicycle work? ; how does a steel ship float?

LIGHT INTERESTS.—Why do glasses magnify? ; how does a camera work? ; how does a cinematograph work? ; why is the sky blue? ; how does television work?

ASTRONOMY INTERESTS.—What are stars? ; about the sun.

HEAT INTERESTS.—How is rain caused? ; how does water freeze? ; how does steam work an engine?

GENERAL SCIENCE INTERESTS.—How does a motor car work? ; about science and scientists.

GEOLOGY INTERESTS.—Cause of earthquakes? ; how was the earth made?

SOUND INTERESTS.—How does the gramophone work? ; how does the piano work?

MAGNETISM INTERESTS.—Why does a compass point north?

CLASS 2.—INTERESTS CONFINED TO A PARTICULAR ENVIRONMENT.

C.=City ; T.=Township ; R.=Rural.

BIOLOGY INTERESTS.—*Human body*. How grows (C.).

Animals. How milk gets inside a cow (R. and C.) ; about horses (R. and T.).

ELECTRICITY INTERESTS.—How tramcar works (C.) ; how does a dynamo work? (C. and T.) ; how is electricity made (C. and T.) ; how does an electric bell work? (C.).

124 *Scientific Interests of Senior School Children*

CHEMISTRY INTERESTS.—How do matches work? (C. and T.); how is coal made? (C.); how do fire extinguishers work? (C.).

MECHANICS INTERESTS.—About pneumatic drills (C.); how does a vacuum cleaner work? (C.).

LIGHT INTERESTS.—How do you see in a mirror? (C.); how does a periscope work? (C.).

ASTRONOMY INTERESTS.—What are comets? (C.).

HEAT INTERESTS.—How does a thermometer work? (C.).

GEOLOGY INTERESTS.—How is sand made? (T.); how do we know where coal is? (T.).

SOUND INTERESTS.—How do talkies work? (C.).

CLASS 3.—INTERESTS CONFINED TO A PARTICULAR SEX.

B=Boy; G.=Girl.

BIOLOGY INTERESTS.—*Human body*. How do teeth grow? (G.); how born? (G.); where has man come from? (B.); what is the difference between a boy and a girl? (G.).

ELECTRICITY INTERESTS.—About Morse code (B.); how an electric train works? (B.); how does an electric light work? (B.).

CHEMISTRY INTERESTS.—How is dynamite made? (B.); how is gunpowder made? (B.); what is petrol? (B.); how is tar made? (G.); about gases (B.); about acids (B.); how are diamonds made? (G.).

MECHANICS.—How do the sun and moon keep up? (B.); how does a submarine work? (B.); what keeps an airship up? (B.).

ASTRONOMY INTERESTS.—Any people on Mars? (B.).

HEAT INTERESTS.—How does a turbine work? (B.).

CLASS 4.—INTERESTS CHARACTERISTIC OF AN AGE.

(NOTE: These questions occurred in only the single age-group indicated by the figures.)

SECTION A.—BOYS.

BIOLOGY INTERESTS (11+).—How does a cat meow?; why has a horse four legs?; why has a camel a hump?; why do leaves fall in winter?; how do plants live in the desert?; how does the oil get in orange skin?; how does the milk get in cocoanuts?

(12+).—How does the head ache?; why has a caterpillar so many legs?

(13+).—How do we taste?; why do we yawn?; how does the kangaroo jump?; how do flowers grow?; why are grass and leaves green?; how does the mould get on cheese?

ELECTRICITY INTERESTS (13+).—About accumulators ; how do wires fuse ? ; how does electricity work ? ; how does a motor horn work ? ; how does an electric fan work ? ; how does a transformer work ?

CHEMISTRY INTERESTS (11+).—Who invented chemicals ? ; how does coal gas burn ?

(12+).—What is steel ? ; how does a film work ?

(13+).—Who invented ink ? ; how is glass coloured ? ; about Neon ; how is the atom split ?

MECHANICS INTERESTS (12+).—Why does blotter suck in ink ? ; how does a fountain pen work ?

(13+).—How has the earth been weighed ? ; how does a buoy float ? ; why are we thrown forward when a train stops ? ; why does a bus not overturn when cornering ?

LIGHT INTERESTS (11+).—How are colours formed ? ; how is light made ?

(12+).—How are mirrors made ?

(13+).—How does the sextant work ? ; how are there such colours in a diamond ? ; what is light ?

ASTRONOMY INTERESTS (13+).—What happens when a comet falls ? ; how high is the sky ?

HEAT INTERESTS (11+).—How radiators work ? ; how is asbestos fireproof ?

(13+).—How is fog caused ? ; how does a thermos flask work ?

GEOLOGY INTERESTS (11+).—How are ores formed ? ; how is gold obtained ?

(13+).—What is a geyser ?

SOUND INTERESTS (11+).—How does the concertina work ?

(12+).—How does a bell work ?

(13+).—How does a dictaphone work ?

MAGNETISM (13+).—How does an electromagnet work ? ; how do we know the world is a magnet ?

SECTION B.—GIRLS.

BIOLOGY INTERESTS (11+).—How the head aches ? ; why has a horse four legs ?

(13+).—How the body works ? ; how is the butterfly formed ? ; how does a glow-worm glow ? ; why cannot amœba be killed ? ; why are leaves, etc. green ?

ELECTRICITY INTERESTS (11+).—How does electricity not pass rubber ? ; how does the switch put on electricity ?

(12+).—How do X-rays work ? ; how do lightning conductors work ?

(13+).—How many volts of electricity kill ?

CHEMISTRY INTERESTS (11+).—Why is some water hard? ; how does coal gas burn? ; how do gas masks work? ; how does salt form colours in the fire? ; how does sugar make the fire flare up?

(12+).—How is iron made?

(13+).—How is poison gas made?

MECHANICS INTERESTS (11+).—How do wheels work?

(12+).—What is gravity?

LIGHT INTERESTS (12+).—How is light made?

(13+).—How are colours formed? ; how are there such colours in diamonds?

ASTRONOMY INTERESTS (11+).—What is our distance from the sun?

(13+).—About the planets.

HEAT INTERESTS (11+).—How does salt thaw ice? ; how does hot water crack thick glass? ; why are the tops of mountains cold when they are nearer the sun?

(12+).—How are ocean currents formed?

GEOLOGY INTERESTS (11+).—What is lava? ; why is the earth round? ; how is china made?

(13+).—How is tin got from the ground? ; what is the inside of the earth like?

SOUND INTERESTS (12+).—How does the violin work?

TOTAL NUMBER OF MAIN TOPICS—170.

The scientific interests indicated by the records of questions children want answering were found to represent the major part of their interest in science. The total number of questions asked on the MAIN topics indicated in Classes 1-4 were counted, as was the total number of questions asked on all science topics. Forty-eight per cent of the total questions were confined to the MAIN topics indicated ; the remaining 52 per cent of the questions were scattered over 412 other miscellaneous scientific topics.

VI.—INFLUENCES ON THE SCIENTIFIC INTERESTS OF SENIORS.

Table IV shows the results of an analysis of the data obtained to find the influence of age on the main scientific interests of seniors ; as far as possible the aim was to decide at what age a particular interest originated. The table was obtained by counting the number of topics in each of the three environments that persist through the age-range 11-13+, i.e., such topics were acquired before the beginning of the senior stage ; in addition, were counted in each environment the topics acquired in ages 12+, persisting to 13+ ; finally were included the topics for each environment acquired in the ages 13+.

TABLE IV.—ORIGINATING AGES FOR MAIN INTERESTS IN SCIENCE AT 13+.

Source.	Total Number of Topics relating to Specified Source.						Per cent Main Interest Field relating to Specified Source.					
	City.		Township.		Rural.		City.		Township.		Rural.	
	B.	G.	B.	G.	B.	G.	B.	G.	B.	G.	B.	G.
Acquired before 11+, persisting to 13+	104	87	35	21	53	38	90	90	34	30	56	46
Acquired in 12+ persisting to 13+	5	3	48	41	21	11	4	3	47	58	22	13
Acquired in 13+	6	7	18	8	20	34	5	7	18	11	22	41
TOTAL MAIN INTEREST FIELD	115	97	101	70	94	83						

B. = Boys; G. = Girls.

Inferences that have been made are :

- (1) Environment has a much greater influence than sex on the age at which interest originates in the main topics arousing the scientific curiosity of senior children.
- (2) City children seem to acquire the bulk of their main interests in science before the 11+ stage, whereas the township children seem to do so in the 12+ stage. Rural children seem to be developing more evenly over the whole age-range up to 13+ in regard to their main interests in science.

Table V gives the results of an analysis to find the proportion of main topics in science that are typical of environment. Topics were taken as typical of an environment where questions were asked only in one of the three types of environment used, or where one environment produced a number of questions that largely outnumbered those on the same topic asked in any other district.

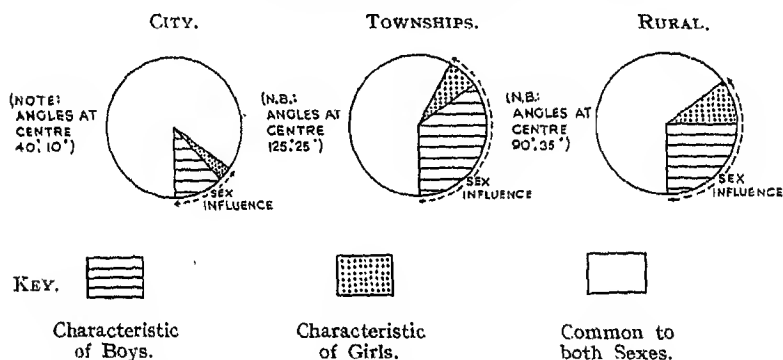
TABLE V.—MAIN TOPICS THAT TYPIFY ENVIRONMENT, 11-13+.

	City.		Township.		Rural.	
	B.	G.	B.	G.	B.	G.
Number Environment typifying Topics	14	16	2	9	3	2
Percentage of the Total Topics in Main Interest Field at 13+	12	16	2	13	3	2

The inference made is that only a small body of topics typify the environment ; in this respect, city life seems to have the greatest effect on the minds of the children.

To gain an indication of the part played by sex in deciding the main topics of interest to senior children, the main topics about which boys only in each environment asked questions were counted and taken as characteristic of boys' interest in science ; then the main topics, about which girls only in each environment asked questions, were counted and taken as characteristic of girls' interest in science ; finally, the main topics, common to both sexes in each environment, were counted. The pictograms represent the results in a pictorial fashion ; the full circle represents the whole interest-field.

PICTOGRAMS TO SHOW THE INFLUENCE OF SEX ON THE MAIN SCIENTIFIC INTERESTS OF CHILDREN.



The inferences made from the pictograms are :

- (1) The largest proportion of the main topics of interest in science to children are the common intellectual property of both sexes.
- (2) Boys seem to have the stronger bias away from this central common core of interests.
- (3) Sex influence on the main scientific interests, though small, seems to be greatest in the townships, closely followed by the rural environment ; in the city group, this influence seems to be much less evident.

VII.—SUMMARY OF CONCLUSIONS.

(1) Boys concentrate their interests largely in the scientific field ; girls' interests are much more widespread with a definite core of scientific interests.

(2) Biology, chemistry, mechanics, and electricity dominate the scientific interests of both sexes in the senior stage.

(3) Boys and girls are interested in the same sciences to rather similar degrees; also, the largest proportion of main topics of interest in science to children are the common intellectual property of both sexes; boys have the stronger bias away from the central common core.

(4) Sex influence on the main scientific interests, though small, seems to be greatest in the township and rural groups; it is least evident in the city group.

(5) A small proportion of the main topics interesting to senior children are characteristic of an age up to 12+.

(6) Environment has a definite influence on the distribution of boys' interests between the non-scientific and scientific fields; this does not apply to girls.

(7) The character of the kind of science in which senior children are interested is somewhat dependent on environment, particularly with rural children where there is a more concentrated interest in a fewer branches of science, compared with the more widespread interest shown in a larger number of sciences by city and township children.

(8) City children acquire the bulk of their main interests in science before the 11+ stage; township children in the 12+ stage; rural children develop their scientific interests more evenly over the whole age-range to 13+.

(9) Only a small body of topics can be taken as typical of environment. City life has the greatest effect in this respect.

Finally, I express thanks to Mr. Drew, Inspector of Schools for the City and County of Newcastle-upon-Tyne Education Authority; to Mr. Paget, H.M. Inspector for the Northumberland and Tyneside Area; and to Mr. Spink, Director of Education for Northumberland, for assistance in collecting data in the various types of environment used.

VIII.—REFERENCES.

- ¹ *Forum of Education*, February, 1928, p. 22.
- ² *Forum of Education*, February, 1928, p. 32.
- ³ *School Science Review*, Nos. 67 and 68, 1936.
- ⁴ *Common Science*: Published Bell.
- ⁵ *British Journal of Educational Psychology*, Vol. 6, 1936, p. 151.
- ⁶ *British Journal of Educational Psychology*, Vol. 3, 1933, p. 18.
- ⁷ *School Science Review*, Nos. 67 and 68; see pp. 571, 322, 323, 565, 568.
- ⁸ *Forum of Education*, February, 1928, p. 30.
- ⁹ *School Science Review*, Nos. 67 and 68; see pp. 322, 323, 565, 568.
- ¹⁰ *Common Science* (Preface, p. v): Washburne (Bell).
- ¹¹ *Adjusting the School to the Child*: Washburne (World Book Co., New York). See p. 111.
- ¹² *Handbook of Suggestions for Teachers* (H.M. Stationery Office, 1937), p. 486.
- ¹³ *Organisation and Curriculum of Sixth Forms in Secondary Schools* (H.M. Stationery Office, 1938), p. 56.
- ¹⁴ *Teaching of General Science* (Science Masters' Association): John Murray, 1938, p. 18, p. 19, p. 28.

RÉSUMÉ.

LES INTÉRÊTS SCIENTIFIQUES DES ÉLÈVES DES ÉCOLES PRIMAIRES SUPÉRIEURES.

L'on recueillit et analysa des feuilles, sur lesquelles les élèves avaient écrit des questions sur des sujets sur lesquels ils voulaient se renseigner. Ces feuilles provenaient d'Écoles Primaires supérieures dans une grande ville, dans des communes, et à la campagne. L'on résume les 170 thèmes principaux de l'intérêt scientifique qui émergèrent des données ainsi rassemblées.

Le sexe exerce une influence puissante à déterminer la distribution des intérêts entre les sciences et les autres sujets, et pourtant il n'a qu'une faible influence sur les intérêts principaux en dedans du champ scientifique. Les thèmes principaux de l'intérêt scientifique appartiennent également aux deux sexes.

Le milieu a une influence non seulement sur le choix des garçons entre la science et le reste mais aussi sur le choix de la branche de la science à laquelle l'enfant s'intéresse, et pourtant il ne s'associe qu'à un petit groupe de thèmes typiques. L'âge auquel les principaux intérêts scientifiques se manifestent est également déterminé par le milieu.

Il y a relativement peu de sujets qui sont caractéristiques d'un certain âge. La biologie, la chimie, la mécanique et l'électricité sont les intérêts scientifiques les plus importants chez les deux sexes.

ZUSAMMENFASSUNG.

DIE WISSENSCHAFTLICHEN INTERESSEN VON SCHÜLERN AN SENIORSCHULEN (DER OBERN HÄLFTE DER ENGLISCHEN VOLKSSCHULE).

Kinder aus grosstädtischen, städtischen, und ländlichen Gegenden wurden aufgefordert, Fragen über Dinge zu schreiben, die sie wissen wollten.

Die Knaben hatten zum grössten Teil Interesse für wissenschaftliche Angelegenheiten; die Interessen der Mädchen waren verschiedenartiger. Doch übte das Geschlecht wenig Einfluss auf die Hauptinteressen, die auf wissenschaftlichem Gebiet lagen, aus. Die Hauptthemen wissenschaftlichen Interesses sind der gemeinsame geistige Besitz beider Geschlechter.

Die Umgebung beeinflusst nicht nur die Verteilung der Interessen der Knaben zwischen Nichtwissenschaft und Wissenschaft; sie beeinflusst auch die Zweige der Wissenschaft, wofür die Kinder Interesse haben; und doch ist sie mit nur einer geringen Zahl typischer Gegenstände verbunden. Das Alter, in dem die hauptsächlichsten wissenschaftlichen Interessen auftreten, wird auch teilweise durch die Umgebung bedingt.

Verhältnismässig wenige Themen sind charakteristisch für ein bestimmtes Alter.

A STUDY OF FACTORS USEFUL IN CHOOSING CANDIDATES FOR THE TEACHING PROFESSION.

BY J. A. LAWTON.

- I.—*The problem. The selection of candidates for the teaching profession without any pre-college score of aptitude for teaching.*
- II.—*The criterion of teaching ability.*
- III.—*The connection between the criterion and*
 - (i) *academic ability ;*
 - (ii) *the non-academic school record ;*
 - (iii) *interview impressions ;*
 - (iv) *a score combining these three factors.*
- IV.—*Choice motives and success at teaching.*
- V.—*Conclusions.*
- VI.—*Bibliography.*

I.—THE PROBLEM.

WITH the passing of the student-teacher year and of the pupil-teacher system it has become almost impossible to obtain a pre-college estimate of aptitude for teaching, based on actual work in the classroom, on which to frame a forecast of the ultimate success or failure of any candidate for the profession. It follows therefore that those responsible for selecting students for the training colleges have to make their judgment on the basis of performances and attributes which are, at best, only indirectly related to teaching ability.

It must be admitted that a number of students, possibly only a small minority, do gain admission to the colleges and afterwards show themselves, for temperamental or other reasons, to be unfitted for the profession they have chosen. It may be perhaps that this minority is no larger than was the case when the criteria for admission to the colleges included a direct assessment of the teaching ability of the applicants. On the other hand the fact that some future failures are admitted leads us to enquire whether the means at present employed for discriminating between successful and unsuccessful candidates are sound or unsound.

The usual standards by which the fitness of any applicant for the profession is judged are (i) his performance at a school-leaving examination, (ii) a record furnished by the head master of his school, (iii) impressions of him formed at an interview, which because of the numbers

of those pressing for admittance, must necessarily be brief. The occasional failure of these standards may be due to the fact that there is little relation between them and the factors which make for success at teaching, or it may be that the best use is not being made of them, or it may be that the fault lies neither in the standards nor in their use, but in the fact that there is an insufficient number of those whom we would wish to see as teachers amongst the original applicant group.

The problem of gauging the effectiveness of these standards is of great importance even if the number of unsuitable persons admitted to training colleges is very small. The entry into college of those who are, for one reason or another, unfitted to teach is a matter which no one can regard with equanimity. Its importance lies chiefly in the fact that unfitness here concerns, not so much the maladjusted person himself, but the great number of children who must inevitably suffer by being placed in his charge. It is a matter of urgent necessity that anyone likely to be unfitted to teach, from whatever cause, should be prevented by every available means from entering the profession. To be fair to him, and to those who are sponsoring his career, it is imperative that he should be stayed prior to his entry into college. In facts such as these lie the importance of calling into question at frequent intervals the methods adopted for the selection of students for the training colleges.

II.—THE CRITERION OF TEACHING EFFICIENCY.

It is necessary first of all to define the standard by which fitness for the profession is to be judged. In the investigation which is to be described all the groups concerned were composed of two-year training college students and the criterion of teaching efficiency was taken to be the teaching grade awarded by the colleges to which they belonged. This mark is open to criticism on the ground that it is awarded at a stage when the person concerned is not mature and that it in no way indicates his ultimate achievement. It has been criticised too, on the score of vagueness by Professor Godfrey Thomson¹³ and others, and as representing the opinion of a small group of persons (the staff of one particular training college) by Cattell⁵. In spite of such criticisms it is still the method in common use for judging the efficiency of training college students and there is a measure of agreement amongst most training college staffs that it furnishes a reasonably reliable indication of the type of teacher the student will ultimately become. Pinsent¹¹ for example, states that "the grades used in the professional year are used largely as a basis for appointment and our experience has been that in the long run, opinions

of teaching capacity based on the grades achieved in the professional year do serve as a guide to future teaching capacity." In addition it is felt that a teaching mark, far from representing the opinion of the staff of one college only, has a value about which there is a considerable measure of agreement amongst inspectors, supervisors, head masters of practising schools and all concerned with the assessment of teaching ability.

Some slight evidence that the college grade has a certain amount of predictive value can be drawn from the following observations. A group of 67 men students were admitted to a two-year training college in the September of 1935 and were given three periods of school practice, each of one month's duration, held in December, 1935, June, 1936, and March, 1937. Only on very rare occasions did any student work under the same supervisor twice in his career, and supervisors did not know the marks which had been awarded to their students on previous practices. It is unlikely therefore that there was "halo" from one practice to the next. In addition a deliberate attempt was made to vary the conditions under which the students worked during the practices, e.g., if a student was given a junior class on one occasion he would be given a senior class on the next. The product-moment correlations between the marks given on the practices were :

<i>1st and 2nd</i>	<i>2nd and 3rd.</i>	<i>1st and 3rd.</i>
0.61 \pm 0.08	0.75 \pm 0.05	0.72 \pm 0.06

More evidence than this is necessary before any far-reaching generalisation can be made, but it does appear that the relative abilities of this particular group of students did not alter much during the period they were at college, those who were above average on entry tending to be above average on leaving, apart from any general improvement which has been effected throughout the entire body of students during their college career. It may be justifiable to assume, although it must be admitted that the evidence is somewhat meagre, that, since the relative abilities of the students tended to remain constant in college, there would also be a little change either prior to admittance or after leaving college ; in other words the better-than-average man would always be better than average at whatever point in his career the measurement was made. If this is the case then a teaching mark awarded before entry to college would be the best criterion for admittance, and the college teaching mark the best guide to future teaching ability in spite of possible exceptions to the rule.

III.—(i) THE CONNECTION BETWEEN THE CRITERION AND
ACADEMIC ABILITY.

It does not by any means follow that a brilliant student will make a good teacher. There have been many attempts in the past to trace a relationship between the factors which make for success in examinations and those which make for success in teaching. Knight⁸ found a correlation of 0.33 between normal school success and general teaching ability. Cooper⁹ with respect to 107 students found a mean square contingency coefficient of 0.332 between marks awarded for college examinations and those for teaching practice. Payne¹⁰ said that all failures in teaching in a group which he examined emanated from the lowest third of classes graduating from High School. Somers¹² found a correlation of 0.707 between the results of a college examination and an assessment of teaching ability in the first year after leaving college. More recently Pinsent¹¹ stated that he failed to find any significant connection between academic record or the abilities measured by intelligence tests and teaching grade for a group of male university students, and Turnbull¹⁴ found small contingency coefficients in groups of university students. All of these results relate to small group, of about a hundred students.

It was my good fortune to have access for a period of years, to the records of students from a group of two-year training colleges, which enabled correlations to be obtained for much larger groups than these. The examination papers were marked first of all by internal examiners and were then checked and re-assessed by external examiners common to the entire group. The teaching marks on an A to E scale were awarded first of all by the training college supervisors and then checked throughout the colleges by a group of visiting inspectors. The teaching classification provides for eleven grades. The correlation coefficients between the two sets of marks were obtained for three annual groups of students.

<i>Year.</i>	<i>No. of Students.</i>	<i>"r."</i>	<i>S.E.</i>
1932	705	0.484	0.029
1936	528	0.447	0.035
1937	497	0.457	0.036

The first significant feature which emerges from a consideration of these results is the high degree of constancy maintained from year to year. This appears to be evidence that there is considerable agreement between inspectors and supervisors with regard to the meaning and value of the marks awarded.

It is evident that there is a connection between the results of examinations and an assessment of teaching efficiency, and it is interesting to speculate on the cause of this. It is unlikely that there is pronounced "halo," good students being marked as good teachers, for whilst supervisors may in some cases know of a student's college record and be influenced by it, the inspector who is finally responsible for the award of the mark can have no such knowledge. Perhaps general intelligence is at least the part cause, although other qualities, perseverance for example, may exert an additional effect, and the actual knowledge required for success in examinations is likely to be of use in teaching.

An additional teaching mark.—The college records contained, in addition to the general teaching mark already mentioned, another mark awarded for the ability to teach physical training. It was considered worth while to obtain the correlations between the latter set of marks and the general examination marks.

Year.	"r,"	S.E.
1932	0.352	0.033
1936	0.324	0.039
1937	0.300	0.036

The correlations are not so high as those obtained in the former case. This is perhaps to be expected, for, whilst it is reasonable to assume that academic knowledge and the qualities required for the acquiring of it may make a person into a better teacher of the academic subjects of the curriculum, it does not seem at all probable that these would make him into a better teacher of a non-academic subject like physical training. It may be wondered why there is any significant correlation at all, but it is very likely that there is a connection between the two sets of teaching marks; in other words a man who is able to teach the general subjects of the curriculum well will be able usually to make a reasonable success of the teaching of any single subject. The actual correlations between the two sets of teaching marks were :

Year.	"r,"	S.E.
1932	0.546	0.026
1936	0.544	0.031
1937	0.500	0.034

in the brief time available an accurate prediction of teaching ability can be made. The experiment described was performed in order to obtain a rough idea of the reliability of the interview in making such an estimate.

Groups of men students from each of three hostels were interviewed by a person experienced in the assessment of teaching efficiency. The teaching mark of each individual in each group was estimated on the usual A to E scale on the basis of an interview lasting not longer than five minutes. These marks were then correlated with marks actually obtained by the students on a teaching practice.

<i>Group.</i>	<i>"r."</i>	<i>S.E.</i>
A	0.56	0.155
B	0.51	0.166
C	0.58	0.160

The same groups were then assessed by the tutors in charge of the hostels, again all men experienced in estimating the ability of students on teaching practice. This time, however, the judgments were based, not on the conclusions drawn from a single brief interview, but on intimate knowledge extending over a period of at least a year. It happened that in one or two cases the tutor was called upon to estimate the mark of a person he had actually seen at work in the classroom. When such cases did occur they were of course omitted from the groups. The correlations between the tutors' estimates and the actual teaching marks were then obtained.

<i>Group.</i>	<i>"r."</i>	<i>S.E.</i>
A	0.80	0.100
B	0.81	0.113
C	0.67	0.145

It appears that the judgment of the first interviewer was no more reliable than if the students had been assessed solely upon the results of an examination. On the other hand the estimates made by the hostel tutors seem to indicate that a person experienced in the estimation of teaching ability should have little difficulty in making a reasonably accurate forecast of the chances of success of any applicant purely on the basis

of an intimate knowledge of him. Head masters of secondary schools should have then a good idea of the possibilities of the boys from their schools who wish to enter the training colleges. The success to be expected of forecasts based on intimate knowledge is probably due to the fact that not only is the assessor in possession of estimates of personality traits, which are all that can be expected to emerge from a brief interview, but that he can also use in formulating his final judgment his knowledge of the intellectual ability, the interests and the character qualities of the persons who are to be assessed.

(iv) *A score combining the three factors.*—Burt⁴ has shown that certain factors can be estimated with considerable success by means of the interview whilst the estimation of others is attended with considerable difficulty. If success is to be obtained in the prediction of teaching ability by means of the interview, it would perhaps be better to limit the use of the method to the estimation of traits which experiment has shown are capable of being so assessed and which are also considered to be requisite to success at teaching. There is little difficulty in deciding what traits are required by the successful teacher. Cattell⁵ has made a survey of professional opinion as to the qualities to be desired and has evolved a tentative scale "whereby head masters and others might make reports upon the prospective value of pupils intending to become teachers." In America there is a wealth of literature upon the subject, Cattell giving an extensive bibliography. In Germany, Keilhacker⁷ has attempted to discover the qualities which the children themselves would desire in their teachers. There is considerable agreement amongst all these workers as to the particular traits required for success, the lists of qualities generally including leadership, kindness, sense of humour, sympathy, tact, enthusiasm, perseverance, and self-control. Of these, perhaps judgments could be framed at an interview of leadership, sense of humour, tact, self-control and maybe enthusiasm; the remainder could only be estimated as a result of intimate knowledge.

But whilst the interview will provide estimates of certain necessary traits it has been shown to be desirable also to take into account the academic ability and the non-academic interests of the applicants, and if possible estimates of character qualities based on long acquaintance. A combined score of this type might provide a very useful guide to future teaching ability.

An attempt to obtain and to estimate the value of such a score was made in the cases of two groups of men students. A score sheet was drawn up in which the individuals were marked with respect to

several attributes considered to be necessary for teaching success. The totals assigned to each trait or performance were, it is to be feared, arrived at quite arbitrarily. The form had four main headings, the first, academic ability, was scored according to the performances of the students in a college examination. The other three were (i) traits which could be estimated at an interview, (ii) interests in non-academic pursuits, and (iii) estimates of character qualities which could only be made on the basis of fairly intimate acquaintance. The first of these was to be estimated by an interviewer after a session lasting not longer than five minutes, the second was completed by a small committee of students, and the third by the hostel tutor. The student committee did not mind estimating their fellows with respect to width of non-academic interests and activities when they were assured of the nature and purpose of the experiment.

THE COMPLETE FORM.

Name.....

<i>Factor.</i>	<i>Assessed by</i>	<i>Maximum Mark.</i>	<i>Mark Obtained.</i>
I. Academic Attainment	Examination	25	
II. Voice	Interview	5	
Appearance		5	
Lucidity of Expression		5	
Brightness		5	
Interest in Affairs		5	
III. Interest in Games	Student Group	5	
Ability at Games		5	
Interest in Social Affairs		5	
Leadership Qualities		10	
IV. Sense of Humour	Hostel Tutor	5	
Kindness		5	
Tact		5	
Industry		5	
Determination		5	
Total Score		100	

It is evident from the form that the task of the interviewer has been made as light as possible, the estimation of traits such as sense of humour and tact, which might conceivably have been assessed by him, being delegated to the hostel tutor. This was done purposely for there would appear to be little chance of the interviewer making a mistake upon the five qualities which he is called upon to assess here.

The total scores obtained upon the form were correlated with the actual teaching marks awarded on school practice, the standard errors in this case being obtained by Fisher's formula.

<i>Group.</i>	<i>No. of Students.</i>	<i>" r."</i>	<i>S.E.</i>
A	45	0.70	0.076
B	45	0.82	0.049

The results appear at least to encourage the belief that could a form of this nature be completed with the same reliability prior to the admission of students to the training college as was possible afterwards, it would afford a fairly reliable means of predicting ultimate success or failure as a teacher. It must be pointed out, however, that the difficulties which would confront such an attempt are at present almost insurmountable. The diversity of school leaving examinations, for example, makes it exceedingly difficult to score the candidates relatively as regards academic attainment, whilst to attempt a relative assessment of the qualities listed in sections three and four amongst youths coming from a large number of schools, would be an almost impossible task.

IV.—CHOICE-MOTIVES AND SUCCESS AT TEACHING.

There is one further question which merits attention, that of the motives which have impelled a candidate to take up teaching as a career. It is to be feared that there is in two-year training colleges, as Professor Valentine¹⁵ and Mrs. Austin¹ have shown to be the case in University Education Departments, a 'substantial minority' who take up teaching from wrong or inadequate motives. Applicants for admission to college who may have elected to follow teaching as a career on account of the amenities it offers in the way of short hours, long holidays or regular wages are unlikely to make a success of the work to the same extent as those who have chosen it from sounder motives.

It may happen that a man possesses all the attributes which should make him into a successful teacher, but such success cannot be expected to follow automatically if his heart is not in his work. The selector therefore is faced not only with the problem of determining the suitability of a candidate, but with that of estimating the soundness of the motives which have impelled him to choose teaching as a career, a task which is in itself fraught with considerable difficulty.

V.—CONCLUSIONS.

It appears, from the results obtained, that there is a considerable measure of connection between the standards usually adopted for determining the fitness of any candidate for admission to the colleges and the teaching mark awarded at the end of the college career (which is regarded as a fairly reliable indication of future teaching ability). It must be remembered, too, that in all the cases which have been considered here, groups have been dealt with which have already been selected with respect to the standards. It is to be expected then, that correlations between these and a measure indicative of aptitude for teaching would have been considerably higher amongst the original applicant groups than was the case amongst the selected student groups. This appears to lead to the conclusion that the normal method of selection should provide a hurdle which it would be difficult for the unfitted candidate to negotiate.

That occasionally such unfitted persons do manage to effect an entry may conceivably be due to one of the following causes. In the first place it is exceedingly difficult to compare the applicants accurately with respect to the standards prior to their admission to college. That factor most easy to assess, namely academic ability, must be judged not on the basis of a single examination but on the results of a number sponsored by many different examining boards. It is almost impossible to compare candidates with respect to non-academic records and out-of-school interests, coming as they do from a large number of schools of widely differing types. It is very possible, too, that the best possible use is not at present being made of the interview. In the second place it is conceivable that there is amongst the original applicant group an insufficient number of those whom we would wish to see as teachers; in which case some colleges might be compelled to admit applicants who are nearer the border-line of success and failure than they would wish. Finally there is the difficulty of elucidating the motives which have impelled the applicant to choose teaching as a career.

The discontinuance of any form of pre-college testing period of aptitude for teaching was long urged on the grounds that such a preliminary estimate was made at too early a stage in the candidate's career to be of any real value. There is small reason to doubt that the systems then in use were open to many abuses, but there was little real evidence to show that estimates framed on such a testing period, if conducted under proper conditions, would have been as valueless as some would have had us believe. The high correlations between marks awarded on successive practices throughout the college career leads to the conclusion that an

estimate of teaching ability based on even a brief pre-college testing period conducted under the usual training college school-practice conditions might furnish the best single criterion on which to judge applicants for admission to the colleges.

The writer wishes to express his appreciation of the help and advice given to him by Dr. Wynn Jones throughout the whole of the enquiry.

BIBLIOGRAPHY.

- ¹ AUSTIN, F. M. : "An Analysis of the Motives of Adolescents for the Choice of the Teaching Profession."—*Brit. Journ. Educ. Psych.*, Vol. I, Pt. 1, Feb., 1931.
- ² BIRD, G. E. : "Pupils' Estimates of Teachers."—*Journal of Educ. Psych.*, 1917, viii, 35.
- ³ BOOK, W. F. : "The High School Teacher from the Pupils' Point of View."—*Ped. Sem.*, 1905, xii, 488.
- ⁴ BURT, C., AND OTHERS : "A Study in Vocational Guidance."
- ⁵ CATTELL, R. B. : "The Assessment of Teaching Ability."—*Brit. Journ. Educ. Psych.*, Vol. I, Pt. 1, Feb., 1931.
- ⁶ COOPER, H. E. : "Correlation of Grades Achieved in Teaching Practice with the Thurstone Group Intelligence Test."—*Ped. Sem.*, 1924, xxx, 176.
- ⁷ KEILHACKER, M. W. : "The Preferences of German Pupils as to the Age, Sex and Appearance of their Teachers."—*Brit. Journ. Educ. Psych.*, Vol. 2, Pt. 1, Feb., 1931.
- ⁸ KNIGHT, F. B. : "Qualities Related to Success in Teaching in 156 Elementary and High School Teachers in Massachusetts."—*Columbia Contributions to Education*, No. 120.
- ⁹ KRAITZ, H. E. : "Characteristics of the Best Teacher as Recognised by the Children."—*Ped. Sem.*, 1896, iii, 413.
- ¹⁰ PAYNE, E. G. : "Scholarship and Success in Teaching."—*Journ. Educ. Psych.*, 1918, ix, 217.
- ¹¹ PRINSENT, A. : "Pre-College Teaching Experience and Other Factors in the Teaching Success of University Students."—*Brit. J. Educ. Psych.*, Vol. 3, Pts. 2 and 3.
- ¹² SOMERS, G. T. : "Pedagogical Prognosis."—*Columbia Contrib. to Educ.*, No. 140, 1923.
- ¹³ THOMSON, G. : "A Rating Scale for Teaching Ability in Students."—*J. Experimental Psych.*, 1921, vi, 75.
- ¹⁴ TURNBULL, G. H. : "The Influence of Previous Teaching Experience on Results obtained by Students in a University Department of Education."—*Brit. J. Educ. Psych.*, Vol. IV, Pt. 1, Feb., 1934.
- ¹⁵ VALENTINE, C. W. : "An Enquiry as to the Choice of the Teaching Profession by University Students."—*Brit. Journ. Educ. Psych.*, Vol. IV, Pt. 3, Nov., 1934.

RÉSUMÉ.

UNE ÉTUDE DES FACTEURS UTILES DANS LE SÉLECTIONNEMENT DES CANDIDATS À L'ENSEIGNEMENT.

Avec la disparition de l'année de l'étudiant enseignant, et du système de l'élève-maître il n'est plus possible d'obtenir avant l'entrée à l'école normale, un pronostic de l'aptitude à l'enseignement basé sur l'observation du travail du candidat dans la salle de classe. La probabilité du succès chez un candidat qui veut entrer à une école normale doit être maintenant calculée uniquement sur les

résultats des examens, les notes de lycée et l'impression qu'on en gagne au cours d'une entrevue.

Les résultats des expériences que l'on décrit ici tendent à démontrer qu'il y a une corrélation assez élevée entre les mesures de ces aptitudes et accomplissements et une mesure de l'aptitude à l'enseignement. Ceci mène à la conclusion que les méthodes de sélectionnement des candidats qu'on emploie ordinairement à présent sont assez efficaces, pourvu qu'on puisse obtenir des renseignements valables, et pourvu aussi qu'il y ait un assez grand nombre de candidats acceptables dans le groupe qui se présente à l'admission. En même temps l'on croit que le meilleur critère de l'aptitude à l'enseignement c'est une note donnée pour le travail pratique pendant un court stage pratique dans une école avant l'admission du candidat à l'école normale.

ZUSAMMENFASSUNG.

EINE UNTERSUCHUNG DER FAKTOREN, DIE BRAUCHBAR SIND, WENN MAN BEWERBER FÜR DIE LEHRERLAUFBAHN WÄHLEN WILL.

Nach dem Verschwinden der einjährigen Praxis für Schüler, die noch in der Schule sind, ist es nicht mehr möglich, eine Einschätzung der Lehrfähigkeit vor Aufnahme in eine höhere Schule zu ermitteln, die eine Beobachtung der Praxis des Kandidaten im Klassenzimmer zur Grundlage hat. Die Wahrscheinlichkeit des Erfolgs irgend eines Bewerbers um Aufnahme in ein Lehrerbildungsinstitut soll jetzt lediglich auf Basis von Prüfungsergebnissen, Schulzeugnissen und nach dem persönlichen Eindruck in einer Unterredung geschätzt werden.

Die Ergebnisse der geschilderten Versuche scheinen zu zeigen, dass ein ziemlicher Grad der Korrelation zwischen Masstäben dieser Eigenschaften und Leistungen und einem Masstab, der auf Lehrfähigkeit hinweist, besteht. Dies führt zum Schluss, dass die gegenwärtig allgemein angewandten Methoden um unter Bewerbern zu unterscheiden, ihre Pflicht genügend erfüllen, vorausgesetzt, dass zuverlässige Informationen vorliegen und eine taugliche Anzahl guter Bewerber sich bei der ursprünglichen Bewerber-Gruppe befindet. Gleichzeitig glaubt man, dass der beste einzelne Masstab, um künftige Lehrfähigkeit vorherzusagen, eine Lehrzensur sein würde, die auf eine kurze Praxis vor Aufnahme des Bewerbers in eine Lehrerbildungsanstalt folgen würde.

A STUDY OF THE CAREERS OF PUPILS WHO ENTER A SECONDARY SCHOOL AFTER A SECOND ATTEMPT IN THE ADMISSION EXAMINATIONS.*

By D. M. COOMBES.

(From the Education Department, University of Birmingham.)

- I.—*Introduction and aim of the investigation.*
- II.—*The material.*
- III.—*A qualitative comparison of the positions in the secondary school of over 800 second "shotters" with the first "shotters" directly above and below each of them in the admission examination.*
- IV.—*A quantitative estimation of the yearly positions of the average second "shotter" compared with the average of adjacent first "shotters" at entry to the secondary school.*
- V.—*A comparison between the School Certificate and Matriculation results of neighbouring first and second "shotters" at the entrance to the secondary school.*
- VI.—*Order of merit of second "shotters" at different stages of secondary education relative to other pupils in their entry group.*
- VII.—*Conclusions.*

I.—INTRODUCTION AND AIM OF THE INVESTIGATION.

THE regulations of many educational authorities permit a child to make a second attempt in the admission examination which qualifies for entry to a secondary school. The question of whether a child who has failed at the first attempt in such an examination should be given a second chance has been a controversial topic in recent years, and at the suggestion of Professor Valentine a large-scale survey was made.

The inclusion in the secondary school of pupils admitted after their second attempt in the admission examination virtually means the exclusion of an almost equal number of younger pupils who are making their first attempt. Numerous grounds, such as different rates of mental development, ill-health, and other inevitable human factors have figured in the appeals for a second chance in the admission examination for those

* Based upon a thesis accepted as part qualification for the degree of M A. in Education at the University of Birmingham.

youngsters who have not done justice to themselves at their previous attempt.

It was the aim of this investigation to answer the question, "Does the child who enters the secondary school after his second attempt in the admission examination justify by his career in the secondary school the additional opportunity given him?" If the second chance is granted we can rightly ask if such action is justified when the secondary school careers of these pupils are compared with those who have been successful first time in the admission examination and yet are, on the average, nine months younger. The importance of the problem is shown by the fact that out of 4,000 pupils who sat for the admission examination here studied, 12 per cent were making their second attempt, and that for similar examinations held in years prior to the introduction of a scheme of maintenance allowances, this figure was as high as 20 per cent. As the merits of the second "shotters" (to use an ugly but convenient term) were to be considered in relation to those of the first "shotters," the latter afford the most serviceable standard by which their progress may be gauged. To ensure reliable results, it was decided to adopt two methods of treatment, which may be briefly expressed thus:

- (1) A comparison between the yearly positions in the secondary school and the results of the School Certificate examination of pupils entering the secondary school after their second attempt in the admission examination, and pupils placed immediately above and below them in that examination in which they have made their first attempt. The second "shotters" progress is measured by that of the first "shotters" next to them at entry.
- (2) The whole group at entry to the secondary school is to be split into three divisions of equal size. The top division, or third, will consist of the third of the whole group who are highest in the admission examination, whilst the bottom division will be the lowest third. The progress of each second "shotter" will be recorded by the division in which he is placed during each year of his secondary school career.

To accomplish these surveys, data covering the careers in the secondary school, marks in the admission examination, and the results of the School Certificate and Matriculation examinations, was collected for 1,200 second "shotters" and over 1,500 first "shotters." Owing to irregular movements of some pupils in the secondary school and the different combinations of subjects studied in the higher forms, the careers of only 800 second "shotters" warranted the desired comparative treatment outlined above.

II.—THE MATERIAL.

Printed schedules were sent to the head teachers of secondary schools asking for the following information : date of birth ; date of admission ; average age of entry group ; form entered ; order in form at the end of first to fourth years ; School Certificate result and remarks. The last item included any disturbances such as illness or transfer.

Under the heading " form entered " it was quite common to find that the new entrants were scattered over three forms, which introduced inaccuracies and some difficulties into the calculation of an individual's position relative to the whole group in which he was admitted. Naturally, numbers in any form or group fluctuate from year to year, tending to become smaller towards the higher forms of the school. This involved another difficulty, for positions of a certain pupil in unequal yearly groups are not directly comparable. For many pupils there is " a parting of the ways " after the Remove Form depending partly on the child's inclinations for the sciences or classics. With different rates of progress and promotion, specialisation, leavers and newcomers, the forms and groups will vary annually in both number and composition. In the absence of a common standard it is impossible to base judgment on the unadjusted positions of any pupil. A percentage basis was a partial solution to this difficulty, and accordingly the positions from their first term to fourth year of all pupils were re-cast. Data covering the careers of 1,200 second " shooters " and 1,500 first " shooters " had accumulated ; such figures being considered sufficiently substantial for the scope of the investigation. The average age of the second " shooters " of both sexes at the time of entry to the secondary school in September was found to be eleven years ten months. These pupils are nine to ten months older than the average child entering after the first attempt. It follows that the second " shooters " will also have had a longer preparation in the elementary school, a factor which must be corrected for in marking the papers of the admission examination.

III.—COMPARISON OF THE POSITIONS IN THE SECONDARY SCHOOL OF 800 FIRST AND SECOND " SHOOTERS."

This part of the enquiry aimed at gauging the progress of each pupil admitted after his or her second attempt in the admission examination relative to that of those directly above and below them in this examination, but who were making their first attempt. Thus, the latter serve as a standard by which one can estimate whether the average second " shooter " does better or worse than the average first " shooter " during his school

career. Such an estimate, incorporating both the school and external examinations, may be of a qualitative or quantitative nature.

In this "sandwich" method, the positions of over 800 second "shotter" were tabulated with the corresponding yearly positions of the first "shotter" immediately *above* and *below* each of them in the admission examination. Inspection of the accumulated data readily shows whether the second "shotter" gains or loses at the expense of the neighbouring first "shotter."

Further, if the numerical value of the position is utilised the degree of such gain or loss may be computed. Before dealing with details of the compilation of tables, an actual example illustrating procedure is given.

TABLE I.—SHOWING THE COMPARISON OF A SECOND "SHOTTER" WITH THE FIRST "SHOTTERS" IMMEDIATELY ABOVE AND BELOW IN THE ADMISSION EXAMINATION.

Pupil's Name.	Admission to Secondary School after 1st or 2nd attempt.	Final Order in Group at end of				
		1st Term.	1st Year.	2nd Year.	3rd Year.	4th Year.
D.K.	1st attempt	54	52	40	32	49
L.M.	2nd "	69	44	91	42	79
S.P.	1st "	68	69	79	84	85

Pupil's Name.	School Certificate and Matriculation Result.	Gain or loss of 2nd "Shotter" over 1st "Shotter."				
		1st Term.	1st Year.	2nd Year.	3rd Year.	4th Year.
D.K.	Passed S.C.	-15	+ 8	-51	-10	-30
L.M.	Failed S.C. }	- 1	+25	-12	+42	+ 6
S.P.	Failed S.C. }					

(Names in order of result of Admission Examination.)

As the size of groups varied from year to year and from school to school, all positions were first expressed as a percentage—a simple and convenient unit for comparisons. The second "shotter" (L.M.) is placed between D.K. and S.P., the first "shotter" above and below him in the admission examination. As all three remain in the same group for four years, positions may be directly compared. At a glance it will be seen that L.M. is thirty places lower than D.K. after fourth year examinations, but six places above S.P.

Fourteen schools (six being girls and eight boys) were included, and the results were summarised as shown in Table II below.

Much of the material collected could not be wholly utilised for comparative purposes for the following reasons :

- (a) The positions of some pupils were unknown as they were absent from the yearly examinations.
- (b) Pupils grouped together at entry were either placed in different forms at the onset or pursued a different type of study in the middle and higher forms of the school. Several head teachers expressed the difficulty and inaccuracy accruing from laborious attempts to recast widely differing orders to estimate positions of those in the original admission group.

RESULTS AND COMMENTS.

Table II below is a summary showing the comparison between the yearly positions of second "shooters" and the first "shooters" immediately *above* them in the admission examination.

The letters "G" and "L" refer to the numbers of second "shooters" gaining or losing places respectively at the expense of the neighbouring first "shooters." Thus, in the first line of the figures below, 272 second "shooters" do better in their first term than the first "shooters" above them at entry, whilst 218 do worse.

TABLE II.

Type of School.	1st Term.		1st Year.		2nd Year.		3rd Year.		4th Year.	
	G.	L.	G.	L.	G.	L.	G.	L.	G.	L.
8 Boys' Schools	136	106	220	219	216	210	162	155	51	62
6 Girls' "	136	112	151	153	125	143	102	100	32	35
TOTAL	272	218	371	372	341	353	264	255	83	97

First term results show that the second "shooters" of both sexes score decisively over adjacent first "shooters." Doubtless this may be largely attributed to maturity and the release from the age handicap imposed at the admission examination. After one year the second "shooters" fare, on the whole, almost exactly as well as the first "shooters" with them at the start; 371 of them have improved their relative positions, whilst 372 occupy lower ones. This "balance" does not by itself suffice to prove equality, for the *amount* of change of position

within the group is ignored. A second "shotter" who is one place higher or lower than his neighbouring first "shotter" will, in the above mode of reckoning, count equally with one who has gained or lost fifty places. Detailed consideration of this quantitative aspect is given later in this section. Comparisons with the first "shotters" below produced results very similar to those shown in the above table. For example, the corresponding first year results were in favour of the second "shotters" by the narrow margin between 378 and 367.

In the second and third years there is almost perfect equilibrium for, if the figures be pooled, the ratio of the gains to losses of the second "shotters" are 1206 : 1208. When pupils have settled down in the school, the second "shotters" appear to fare just as well in the middle forms as the younger associates with whom they were coupled on entry. A margin of two in over 1200 supports the earlier suggestion that age, though of aid at the onset, becomes rapidly less influential with time. In their fourth year the second "shotters" of both sexes "lose ground" compared with the first "shotters" above and below them at entry. The advantage of age seniority has now vanished entirely and there remains a small balance in favour of the first "shotter." It will be noticed that in the first year results over 700 second "shotters" were used for the comparison, whereas in the fourth year the number had shrunk to under 200, which would certainly reduce the degree of reliability. The differences between the comparisons of the same second "shotter" with the first "shotters" above and below are as slight as those in the admission examination. Throughout the second "shotters" do slightly better relative to the first "shotters" below than above them, a conclusion which one might have conjectured from their intermediate position.

No deviation from the above conclusions is found if the sexes are segregated, for in both cases the second "shotters" lead at the lower part of the school; just "hold the course" in the middle forms; but are not quite so good in the higher forms and external examinations.

Owing to departures, absences, different courses of study, the number of pupils included in this part of the enquiry dwindled considerably in the very years upon which evidence was particularly desired. Extensive trustworthy information was required for the end of the pupil's school career, and therefore more attention must be given to the academic results in the fourth year and external examinations, for unfortunately they are still by many regarded as the criterion of success and attainment. The final small, but "all round" advantage in favour of the first "shotters" by no means suffices to condemn the calibre of the second "shotters," or to warrant any general modifications in the existing admission examination régime.

IV.—A QUANTITATIVE ESTIMATION OF THE YEARLY POSITIONS OF THE AVERAGE SECOND "SHOTTER" COMPARED WITH THE AVERAGE OF THE ADJACENT FIRST "SHOTTERS" AT THE ENTRY TO THE SECONDARY SCHOOL.

The few paragraphs included under this heading are a supplement to the preceding section inasmuch as the extent of gain or loss of the average second "shotter" relative to the first "shotters" above and below has been investigated. The positions in the group of the pupils to be compared were expressed as a percentage before the average gain or loss of the second "shotter" was computed.

On the right-hand side of Table I, it will be seen that in a group of 100 the second "shotter" (L.M.) is placed lower than D.K., who was the first "shotter" above him in the admission examination. A case in which thirty-one comparisons were made might be quoted as an example. The thirty-one second "shotters" were each placed beside first "shotters" above them on entry. After expressing their positions as a percentage, the sum of the gains between yearly examinations totalled 146, whilst the corresponding losses were 169. The average loss per pupil was, therefore $\frac{169 - 146}{31} = .76$. In the table this is recorded as -1 , which

interpreted, means that in a group of 100 the average second "shotter" will be about one place or 1 per cent lower than that of the first "shotter."

In compiling the tables below, variations naturally occurred within a school as well as from school to school. This forbids value to be attached to a single result. Pooling the results of fourteen schools for four years gives a more acceptable value, for in the summation of 434 averages, extremes will tend to cancel out.

TABLE III.—SHOWING THE AVERAGE PERCENTAGE GAIN OR LOSS IN POSITIONS FOR SECOND "SHOTTERS" IN RELATION TO ADJACENT FIRST "SHOTTERS."

(a) COMPARISON OF SECOND "SHOTTERS" WITH FIRST "SHOTTERS" *above THEM AT THE ADMISSION EXAMINATION.*

	1st Term.	1st Year.	2nd Year.	3rd Year.	4th Year.
Average for Boys and Girls ..	+3.26	+1.44	+ .27	+ .87	-3.63
Average for Boys	+3.42	+ .77	+1.81	- .50	-3.94
Average for Girls	+3.06	+2.52	-1.78	+2.31	-3.09

(b) COMPARISON OF SECOND "SHOTTERS" WITH FIRST "SHOTTERS"
below THEM AT THE ADMISSION EXAMINATION.

	1st Term.	1st Year.	2nd Year.	3rd Year.	4th Year.
Average for Boys and Girls ..	+5.44	+ .14	+ .21	- .31	-2.27
Average for Boys	+6.79	- .76	+ .93	-2.21	-2.33
Average for Girls	+3.73	+1.57	-1.00	+2.89	-2.18

In the above it is seen that the second "shooters" gain progressively diminishes in size and in the fourth year has been converted into a loss equal to the initial advantage. This applies to both sets of comparisons, and generally to both sexes. The second "shooters" naturally do better when compared with the first "shooters" below them than with those above.

Changes of considerable magnitude mark the "upheaval" between entry positions and those of the first term and first year, and contrast with the small gradual transition witnessed in later years. Such an inference has already resulted in the previous section, and clearly shows that the age seniority on the side of the second "shotter" is operative in the first term and first year results. The abnormality in the first term average for the boys in Table IV(b) is partly due to four schools being unable to supply the data and thus materially reducing numbers on which averages were based.

In the fourth year the second "shooters" of both sexes are lower than the corresponding first "shooters" above and below them at the onset. This general backwardness at the close of school days of the average second "shotter" compared with his more youthful fellow entrants is an important conclusion from this analytical study for, in making degree rather than the number of gains or losses the vital factor, the many bare successes of second "shooters" are kept in true proportion and perspective.

An initial advantage, a "neck-to-neck" struggle leading to a slight final "lag" at the end of the race, expresses the academic progress in the secondary school of the average second "shotter" when compared with the average first "shotter" with whom he was coupled on admission.

V.—SCHOOL CERTIFICATE AND MATRICULATION RESULTS OF NEIGHBOURING
FIRST AND SECOND "SHOTTERS."

The School Certificate and Matriculation examinations, being external to the school, serve as an invaluable additional standard by which the merits and demerits of first and second "shooters" may be

judged. Strict comparison is not legitimate here, for many were the numbers and combinations of subjects taken by those to be compared, and though such groupings of subjects would align with the attainments and interests of individual pupils in them, results would certainly differ considerably without a choice of subjects. As with the admission examination, the "luck of the paper," ill health, and other inevitable external factors are still influential.

The important point is what proportion of second "shotter" entrants gain School Certificates in their fourth year as compared with the first "shotters" who were their immediate neighbours in the entrance examination.

Four hundred and fifty-one neighbouring "pairs" of first and second "shotters," who entered the secondary school at the same time, gave the following interesting results, though some of them were not entered for the School Certificate examination.

TABLE IV.

<i>Number and Type of Pupils at entry to the Secondary School.</i>	<i>Number and Percentage gaining School Certificate in their 4th Year.</i>	<i>Number gaining Matriculation.</i>	<i>Number gaining Higher School Certificate.</i>
451 First "Shotters"	221 (49%)	96 (21.2%)	14 (3.1%)
451 Second "	161 (35.7%)	59 (13.1%)	7 (1.5%)

Whilst the School Certificate results only refer to those taking the examination in their fourth year, the Matriculation and Higher School Certificate results naturally include subsequent years. Each of the 451 first "shotters" was ranked next in the entrance examination to a second "shotter" with whom each was compared, though no distinction was made between those first "shotters" who were above and those who were below the neighbouring second "shotter."

The above School Certificate results show the unquestionable superiority of the first "shotters," and if the Matriculation and Higher School Certificate figures are perused, the relative gain is even greater. We know that the second "shotters" are, on the average, nine to ten months older than the average age of the first "shotters," and have also noted a distinct "lag" on the part of the second "shotters" in their fourth year. Such factors partially account for these very conspicuous differences.

As further evidence, the average School Certificate results of 1,136 first "shotters" and 508 second "shotters" who completed a four-year course in the secondary school might be quoted.

TABLE V.

<i>Number and Type of Pupils entering the Secondary School.</i>	<i>Number of these Pupils gaining :</i>		
	<i>School Certificate.</i>	<i>Matriculation.</i>	<i>Higher School Certificate.</i>
1136 First "Shotters".	577 (50·8%)	262 (23%)	51 (4·4%)
508 Second " "	171 (33·6%)	66 (12·9%)	8 (1·5%)

These results, based on considerably larger numbers, though not comparable as their distribution in the entry order to the secondary school would differ, substantiate conclusions made above.

Summarizing, it is noted that the "lag" of the second "shotters" in their fourth year behind the first "shotters" with whom they were coupled on entry, is reflected to an even greater extent in the School Certificate examination. Out of 451 "pairs" of neighbouring entrants, 49 per cent of the first "shotters" have obtained a School Certificate in four years, yet only 35·7 per cent of the second "shotters" have achieved this. In the Matriculation and Higher School Certificate examination the "gulf" is wider still. After School Certificate stage the age seniority will mean that there will be more leavers among second "shotters" than among the first "shotters."

VI.—ORDER OF MERIT OF SECOND "SHOTTERS" AT DIFFERENT STAGES OF SECONDARY EDUCATION RELATIVE TO OTHER PUPILS IN THEIR ENTRY GROUP.

Success of the second "shotter" has been estimated by his performance relative to neighbouring first "shotters" and will now be gauged by using the other members of the group in which he is placed as a standard. Thus essentially the standard by which the second "shotter" is adjudicated has widened from the single first "shotter" to the group of which both are members. The whole admission group is divided into three divisions or compartments of equal size, the top division will consist of the third of the pupils who are highest in the mark list, and the bottom division is composed of the third who are lowest. To obtain a correct perspective of the secondary school careers of second "shotters" three tabulations were necessary:

- (a) Where the pupils were graded according to their order in the admission examination; such order being compared with the respective yearly examination within the secondary school. Variations in the size of the originally equal divisions give an

indication of the movement which takes place between the admission examination and periods in the secondary school, and should substantiate or contradict later findings on the degree of prognostication of the admission examination.

- (b) The pupils are graded from their order in the secondary school after one year; such order being compared with those of later years. The stability within the secondary school, after allowing the pupils a year in which to become "acclimatized" to the new régime and surroundings is here the claimant to our attention.
- (c) Employing the School Certificate and Matriculation examinations as a rough gauge of success in the secondary school, the results of these are classified according to the division into which each of the second "shooters" was placed on entry, and after one year in the secondary school.

This extends (a) by linking the admission and School Certificate examinations between which is the period of secondary education. In addition (b) is further developed by comparing an early school order with that of an external examination taken three years later.

MATERIAL, METHOD, RESULTS AND CONCLUSIONS.

Each second "shooter" was placed on entry to the secondary school into the top, middle or bottom third or division, and the careers of those in each of such divisions has been separately studied.

If the total number in the admission group was 150, then all second "shooters" whose placings were between one and fifty are assigned to the top third, and similarly the bottom third claimed those occupying positions between 100 and 150.

TABLE VI.—THE DIVISION OF THE WHOLE ENTRY GROUP INTO THIRDS, SHOWING THE POSITIONS OF 508 SECOND "SHOOTERS" FROM ADMISSION TO THE FOURTH YEAR IN THE SECONDARY SCHOOL.

<i>Division of Group into Thirds.</i>	<i>Number in each Division after :</i>				
	<i>Entry.</i>	<i>1st Year.</i>	<i>2nd Year.</i>	<i>3rd Year.</i>	<i>4th Year.</i>
Top Third	117	163	156	162	150
Middle Third	167	190	211	193	197
Bottom Third ..	224	155	141	153	161
TOTAL	508	508	508	508	508

It will be seen that of the 508 second "shooters" studied, 117 were in the top third of the entry group, 167 in the middle third, and 224 in the bottom third. That the largest number of second "shooters" at entry to the secondary school were in the bottom third indicated a slight initial disadvantage compared with their fellow entrants.

When the same second "shooters" are recast again relative to the whole of their entry group, but after their first year examinations in the secondary school, they show considerable improvement, the number in the top third having increased by 40 per cent, whilst the decrease in the bottom third is over 30 per cent. Second, third and fourth-year figures show that this advantage does not last, and that the brighter younger pupils steadily gain on the less bright older ones. The defect of the above method of classification is that no indication is given of the movements of individual pupils during their secondary school careers. As an example, it may be asked how are the 117 second "shooters" in the top third of the entry group actually distributed after their first and subsequent years of secondary education. It is only to be expected that some will be relegated to lower divisions or lower positions within the same division. It is equally important to know how many of the 224 in the bottom third at entry will have climbed to the top third after four years, and also the number who continue to occupy the lowest positions.

TABLE VII IS DESIGNED TO SHOW SUCH MOVEMENTS OF THE ABOVE 508 SECOND "SHOOTERS."

<i>Division of Group into Thirds.</i>	<i>Number in each Division classified from Entry Order.</i>				
	<i>Entry.</i>	<i>1st Year.</i>	<i>2nd Year.</i>	<i>3rd Year.</i>	<i>4th Year.</i>
Top Third	117	64 (T.) 35 (M.) 18 (B.)	63 (T.) 41 (M.) 13 (B.)	66 (T.) 37 (M.) 14 (B.)	64 (T.) 32 (M.) 21 (B.)
Middle Third	167	49 (T.) 69 (M.) 49 (B.)	57 (T.) 65 (M.) 45 (B.)	60 (T.) 64 (M.) 43 (B.)	48 (T.) 75 (M.) 44 (B.)
Bottom Third ..	224	50 (T.) 86 (M.) 88 (B.)	36 (T.) 109 (M.) 79 (B.)	36 (T.) 92 (M.) 96 (B.)	38 (T.) 90 (M.) 96 (B.)
WHOLE GROUP TOTAL	508	508	508	508	508

(The letters T, M and B refer to the top, middle and bottom thirds in which the pupils were placed.)

Sixty-four of the 117 pupils assigned to the top third at entry still remain in that category after their first year in the secondary school, the remainder being divided between the middle and bottom thirds. The consistency of the distribution of these better pupils after they have settled down in their new school cannot escape notice. The fourth year figures (64 top third, 32 middle third, and 21 bottom third) is almost identical with the distribution in the first year (64 top third, 35 middle third, and 18 bottom third). Such figures clearly indicate that the majority of second "shooters" who rank highest in the entry order to the secondary school are those who benefit most from the secondary school course.

Of those originally placed in the bottom third, only 16 per cent climb ultimately to the top third, the remainder failing to make any substantial progress relative to other members of their group—a result shortly to be justified by School Certificate results. It would be very unjust to condemn all pupils placed in the lowest third, for inquiry into the vocations and careers of some of these "plodders" shows that some of them make enormous strides after leaving school, and meet with success in the examinations of professional bodies.

If order of entry is the criterion of classification, the second "shooters" have a relative advantage after the first year in the secondary school. This lacks permanence and at the end of school days the second "shooter" is below the standard of the average first "shooter." If the internal order of the pupils in the secondary school after one year be made the basis for classification, the uncertain variables associated with the admission examination and the early but peculiarly evanescent upheaval of the first year, are eliminated. Appropriately this forms the subject matter of the next section.

THE MOVEMENTS OF 508 SECOND "SHOOTERS" IN THE SECONDARY SCHOOL AFTER CLASSIFICATION FROM FIRST-YEAR ORDERS.

It will be agreed that after one year in the secondary school the average child will have become accustomed to the régime of the new environment, and that the first annual examination should be a fairly reliable basis for grading if it is desired to trace the secondary school performance of certain groups of pupils.

In the previous section it has been clearly shown that early fluctuations are definitely evanescent by the close of the first year, which must, therefore, be a surer starting point. The 508 second "shooters" were each allocated to either the top, middle, or bottom thirds as justified by their positions after one year.

TABLE VIII.

Division of Group into Thirds.	Number in each Third, being classified from 1st Year Order in the Secondary School.			
	1st Year.	2nd Year.	3rd Year.	4th Year.
Top Third	163	{ 121 (T.) 39 (M.) 3 (B.)	117 (T.) 37 (M.) 9 (B.)	107 (T.) 42 (M.) 14 (B.)
Middle Third	190	{ 26 (T.) 129 (M.) 35 (B.)	38 (T.) 115 (M.) 37 (B.)	31 (T.) 108 (M.) 51 (B.)
Bottom Third	155	{ 9 (T.) 43 (M.) 103 (B.)	7 (T.) 41 (M.) 107 (B.)	12 (T.) 47 (M.) 96 (B.)
WHOLE GROUP TOTAL	508	508	508	508

Out of the 163 pupils credited to the top third after one year in the secondary school, 65 per cent of them are still there after four years, whilst the corresponding percentage for the bottom third is 62 per cent. As an expression of consistency within the school, these figures are positive, and, when all the disturbing factors are considered, they are relatively high. If the analysis be perused more carefully, it will be seen that from the top third there is a slight yearly increase in the numbers relegated to the middle and bottom thirds, but that this is offset by promotions from the lower thirds. In drawing together material for the table, the cases of pupils who in one year were just in one division or third and outside it next year, only narrowly to return in the following year, were common. Such movement, though actually small and not recorded for those around the median of their division, influences the figures of the table.

The middle third provides an almost equal proportion of members for the top and bottom thirds, the latter receiving a larger number in the fourth year. Both the middle and top thirds show a larger quota for the bottom division in the fourth year. *It appears both from the table and study of individual second "shooters," that these pupils tend to "slow-up" or "tire" after the third year—an effect seen both in the School Certificate examination as well as the fourth yearly school tests.*

SCHOOL CERTIFICATE AND MATRICULATION RESULTS OF 508 SECOND "SHOTTERS" WHO WERE CLASSIFIED FROM THE ADMISSION EXAMINATION AND FIRST-YEAR ORDER IN THE SECONDARY SCHOOL.

The primary aim here is to trace the performance of each second "shotter" in the School Certificate examination relative to the one of the three divisions in which he or she was placed on admission and after one year in the secondary school.

The preceding conception of the three divisions or groups of equal size, in which position or order of merit is the determining factor for constituents, is again employed. The following table summarizes the fortunes in the leaving examination of 508 second "shotters" who were graded into top, middle, or bottom divisions after one year in the secondary school.

TABLE IX.—SCHOOL CERTIFICATE RESULTS FOR SECOND "SHOTTERS" CLASSIFIED FROM FIRST-YEAR SECONDARY SCHOOL ORDER.

<i>Division of Groups into Thirds.</i>	<i>Number of Pupils after one Year.</i>	<i>Number of such Pupils gaining :</i>		
		<i>School Certificate.</i>	<i>Matriculation.</i>	<i>Higher School Certificate.</i>
Top Third	163	93 (=57%)	46	6
Middle Third	190	57 (=30%)	14	1
Bottom Third	155	21 (=14%)	6	1
TOTAL	508	171	66	8

The top third secures 54 per cent of all the School Certificates, and 69 per cent of the total Matriculation Certificates. On the other hand, the bottom third can only claim percentages of twelve and nine respectively. From this and previous researches¹, we know that after one year in the secondary school, pupils are "settling down" and their positions even at that stage bear high correlations with those of later years. *The inequality of distribution shows that the second "shotters" who are highest in the lower forms of the school will form the nucleus of the successes in the external examination at the termination of the school career.*

A single year in the secondary school gives a reliable indication of attainment three years later, and at this early stage many pupils found to be unsuited to one type of secondary education could be profitably transferred to a school offering a more appropriate type.

¹ See *The Reliability of Examinations*, by C. W. Valentine, p. 87.

In two groups of equal size the ratio of passes in the School Certificate examination is no less than 57 to 14, from which it appears that for many in the bottom division after one year in the secondary school, a return to the senior elementary school or technical school, as recommended by the Hadow Committee should, in the interests of the individual pupil and the taxpayer, receive immediate consideration.¹

TABLE X.—SCHOOL CERTIFICATE RESULTS OF SECOND "SHOTTERS" CLASSIFIED FROM ORDER OF ADMISSION TO THE SECONDARY SCHOOL.

Division of Group into Thirds.	Number of Pupils at Admission.	Number of such Pupils gaining :		
		School Certificate.	Matriculation.	Higher School Certificate.
Top Third	117	57 (=49%)	27 (=23%)	5
Middle Third	167	61 (=36%)	19 (=11%)	1
Bottom Third	224	63 (=28%)	20 (=9%)	2
TOTAL	508	171	66	8

Though numerically the School Certificate "spoils" are divided almost equally between the three groups, it will be noted that the bottom third includes almost twice as many second "shooters" as the top one. A truer perspective is obtained when the School Certificate results are expressed as percentages, for this shows that only one in four of those in the bottom third at entry will obtain a certificate, and less than one in ten will matriculate.

However, examinations must not become the "idols" for us to appease, for many of the life values accruing from secondary school education are not recorded in examinations, and aptitude in one direction may be a positive hindrance rather than an asset in satisfying the general requirements of to-day.

VII.—SUMMARY OF RESULTS AND CONCLUSIONS.

(1) As roughly 12 per cent of the candidates sitting for the secondary school admission examination were making their second attempt, it is reasonable to expect that those of them who receive secondary education should justify in the school and external examinations this additional opportunity.

¹See *Report on the Education of the Adolescent* by Consultative Committee appointed by Board of Education, Chapter III, Para. 97; and Chapter VII, Para. 156.

(2) Those who enter at their second attempt are on the average nine months older than those admitted at the first attempt. This seniority in age and the consequent longer period of preparation in the primary school gives the younger first "shooters" an age allowance in the admission examination, but not in the examinations in the secondary school.

(3) The average position in the entrance examination gained by the second "shooters" is far below that of the first "shooters." If the order of merit of all pupils who enter the secondary school is divided into three equal groups, there are nearly twice as many second "shooters" in the bottom group as in the top group.

(4) Information concerning the secondary school careers of 1,200 second "shooters" and 1,500 first "shooters" was collected, but owing to irregular movements of pupils in the secondary school and the different combinations of subjects studied, the careers of only 800 second "shooters" could be used for comparison with the first "shooters" immediately above and below them at the admission examination. At the end of the first term and first year in the secondary school the second "shooters" on the average gained in position. The absence of the age "handicap" imposed in the admission examination, together with the extra nine months in the primary school, will largely account for this.

(5) There is little to choose between first and second "shooters" in the middle forms of the secondary school, though in their last year there is a pronounced "lag" on the part of the second "shooters," and is well shown by an extract from a table:

AVERAGE NUMBER OF PLACES LOST BY THE SECOND "SHOOTERS"
BY THE END OF THE FOURTH YEAR.

(a) Compared with first "shooters" immediately	}	Boys	3.94
above them at admission		Girls	3.09
(b) Compared with first "shooters" immediately	}	Boys	2.33
below them at admission		Girls	2.18

(6) The superiority of the first "shooters" is confirmed when reference is made to external examinations. Out of 451 pairs of neighbouring first and second "shooters" at admission, 49 per cent of the former have secured a School Certificate after four years, but only 35.7 per cent of the second "shooters" have achieved this. The corresponding percentages for Matriculation were 21 and 13.

(7) A comparison between the School Certificate results and the order of merit of the second "shooters" on entry, and after one year

in the secondary school, cannot but arouse thought on the question of the advisability of a review of the progress of each individual in the early stage of secondary school life. When the entry group is divided into thirds it is found that there are almost twice as many second "shooters" in the bottom third as the top one, yet of the latter 49 per cent secure School Certificates compared with 28 per cent for the former. Fifty-seven per cent of the 163 second "shooters" in the top third at the end of one year get School Certificates. The corresponding percentage for the 155 in the bottom third is only 14 per cent.

(8) Generally it may be said that the second "shooters" are slightly below the first "shooters" in the admission examination. They improve their positions in the first year of secondary school life, being freed from the handicap of age allowance. At the end of four years they are about three places lower in class than their neighbours at entry, and successes in the School Certificate and Matriculation examinations are not so numerous. On the other hand it must be pointed out that the results in these examinations for those second "shooters" who were in the top third of the admission list were almost equal to those of pupils who entered the secondary school after their first attempt. For these pupils a second attempt to enter the secondary school is justified by results.

RÉSUMÉ.

UNE ÉTUDE DE LA CARRIÈRE DES ÉLÈVES QUI ENTRENT À UNE ÉCOLE SECONDAIRE APRÈS LEUR SECOND ESSAI À L'EXAMEN D'ADMISSIBILITÉ.

12% de 4,000 élèves se présentant à un Examen d'Admissibilité en étaient à leur second essai. Le but de la recherche était de tracer la carrière de tels élèves à l'École Secondaire et au "School Certificate Examination."

L'on établit une comparaison entre 800 élèves admis à leur second essai et les élèves voisins au-dessus et au-dessous d'eux, mais qui entrèrent à l'École Secondaire après leur premier essai. L'on trouva que ceux-là gagnent au dépens de ceux-ci pendant le premier trimestre et la première année dans l'École Secondaire. Cet avantage était dû en grande partie au fait qu'ils étaient plus âgés de neuf mois que ceux avec lesquels on les comparait; de plus, ils n'étaient pas, dans l'École Secondaire, pénalisés à cause de leur âge comme dans l'Examen d'Admissibilité. A la fin de leur carrière scolaire ils ne réussissaient pas tout à fait aussi bien, et dans les examens du lycée, et dans les externes, que les élèves plus jeunes à qui ils étaient associés à l'entrée.

Les résultats semblent justifier le second essai à l'Examen d'Admissibilité, des élèves incapables découverts après une année dans l'École Secondaire devraient être transférés à un type d'école qui leur conviendrait mieux.

ZUSAMMENFASSUNG.EINE UNTERSUCHUNG DER LAUFBAHNEN VON SCHÜLERN, DIE BEIM
ZWEITEN VERSUCH DER AUFNAHMEPRÜFUNG DIE HÖHERE SCHULE
BEZIEHEN.

12 v.H. von 4,000 Schülern, die eine Aufnahmeprüfung versuchten, machten schon ihren zweiten Versuch. Die Untersuchung sollte die Laufbahnen solcher Schüler in der höheren Schule und in der School-Certificate-Prüfung verfolgen.

Man machte einen Vergleich zwischen 800 Schülern, die nach ihrem zweiten Versuch bei der Aufnahmeprüfung aufgenommen wurden, und denen, die gerade über und unter ihnen in der Prüfung abschnitten, die aber nach ihrem ersten Versuch die höhere Schule bezogen. Man stellte fest, dass die ersteren Vorteile erlangen auf Kosten der letzteren während des ersten Semesters und des ersten Jahres in der höheren Schule. Diese Vorteile sind grösstenteils auf die Tatsache zurückzuführen, dass sie neun Monate älter waren als diejenigen, mit denen sie verglichen wurden; auch in der höheren Schule wurden sie nicht durch verschiedenes Alter behindert wie in der Aufnahmeprüfung.

Am Ende der Schulzeit erzielten sie etwas weniger Erfolg sowohl in der Schule als bei Prüfungen ausserhalb der Schule, als die jüngeren Schüler, mit denen sie zusammen eintraten.

Die Ergebnisse scheinen den zweiten Versuch bei der Aufnahmeprüfung zu rechtfertigen; ausgesprochene "Versager," die nach einem Jahr in der höheren Schule entdeckt werden, sollten einem geeigneten Schultyp zugeführt werden.

THE BEARING OF GENERAL AND SPECIAL ABILITIES UPON SCHOLASTIC SUCCESS AT THE BEGINNING AND END OF A SECONDARY SCHOOL CAREER.*

BY MARY ORMISTON.

PART I: JUNIOR CHILDREN.

- I.—*Introduction: the problem.*
- II.—*The preliminary work.*
 - (i) *The tests used.*
 - (ii) *The analysis of the preliminary results.*
- III.—*The main experiment.*
 - (i) *The Junior Group.*
 - The analysis of the results of the Junior Scholarship examinations in Districts A and B.*
 - The analysis of the tests involving the spatial factors.*

I.—THE PROBLEM.

At the present time the scholarship examination, taken by pupils of ten or eleven years of age, is an important means of deciding fitness for secondary education. After satisfying this first criterion, pupils embark on a four or five years' course, at the end of which they are judged by a second criterion of success such as the School Certificate or some similar examination.

In the hands of well-organised authorities these two criteria may, perhaps, be regarded as the best available measures of success at the beginning and end of the secondary school career.

The aim of this research was to find the bearing of general and special abilities upon these two examinations.

II.—THE PRELIMINARY WORK.

- (i) *The tests used.*

A considerable amount of preliminary work had to be done in order to obtain tests of general and special abilities from which the analysis could be undertaken.

* Based upon a thesis approved for the degree of Ph.D. in the University of Leeds.

Evidence of both general and special factors has been found in tests already published, and some of these tests formed the basis for the first battery. Verbal and non-verbal tests for *g*, tests for a number, verbal, and spatial factor were made on similar lines to those used in previous work, where such factors had been traced.

The following twenty tests were ultimately formed or selected and applied to 100 elementary school children of ten and eleven years of age: Mechanical Arithmetic¹, Non-Mechanical Arithmetic², Verbal Analogies, Verbal Classification³, Logical Selection⁴, Opposites⁵, Verbal Imagery⁶, Finnish Words Test⁷, German Words Test, Non-Verbal Analogies⁸, Non-Verbal Classification⁹, Correlate Education¹⁰, Dot Pattern Perception¹¹, Overlapping Shapes¹², Paper Cutting¹³, Spatial Symbols¹⁴, Paper Formboard¹⁵, Spatial Imagery¹⁶, the Passalong Test¹⁷, and a Tracing Test. In this Tracing Test the children were required to trace two figures. They were told that they would be marked for accuracy *and* for the amount of the figure completed. The figures were divided into sixty sections, and the number of sections completed to the required standard of accuracy was then divided by the number of sections finished, and thus their final score was obtained. These results were correlated with the estimates of 'persistency' given by the teachers ($r = .49$ P.E. .04).

A similar battery of tests was prepared for children of fifteen years of age. In so far as was possible these tests were made comparable to those used for the younger children, the same fundamentals being involved in each test respectively. The time allowed and the difficulty of the items were adjusted to obtain a satisfactory distribution of marks. The Passalong Test was not used with the older children, because of the length of time needed to administer this individual test.

(ii) *The analysis of the preliminary results.*

From this analysis it was hoped to find the factor loadings of the tests used, and ultimately to find tests highly loaded with the factors

¹ and ² Based on the work of COLLAR: *Brit. Jour. Psych.*, 1920; and FOURACRE: *The Forum of Education*, iv. 3, 1926.

³, ⁴, and ⁵ After TOMLINSON's West Riding tests.

⁶ After L. WYNN JONES: *Introduction to the Theory and Practice of Psychology*, p. 58.

⁷ There was some similarity of form between the German and English equivalents, but not between the Finnish and English.

⁸, ⁹, ¹⁰ and ¹¹ After the tests of EL. KOUSEY: *Brit. Jour. Psych.*, Mon. Supp. xx.

¹² After STEPHENSON: *Brit. Jour. of Educ. Psych.*, 1931.

¹³ After BINET's Paper Folding test.

¹⁴ After ORMISTON: *Geography*, vol. xx, pt. iii, 1935.

¹⁵ and ¹⁶ Published by Stoetling and Co., Chicago.

¹⁷ ALEXANDER: *Brit. Jour. Psych.*, 1932.

detected which would be sufficiently pure tests of these same factors as to serve as axis tests for rotations for the respective factors.

The loadings of the general factor in the tests used was first found. The Spearman¹ method was applied for this purpose. The tests of Non-Mechanical Arithmetic, Verbal Analogies, Paper Formboard, and Non-Verbal Classification were together used as reference values, since their intercorrelations with the battery was large, while the overlap as shown in the specific links between the tests was low.

The *g* loadings obtained thus are given below :

	<i>g</i> loading.
(1) Paper Folding420
(2) Passalong204
(3) Correlate Eduction.....	.456
(4) Spatial Imagery667
(5) Verbal Imagery681
(6) Non-Mechanical Arithmetic824
(7) Mechanical Arithmetic482
(8) Finnish Words.....	.239
(9) German Words.....	.198
(10) Opposites726
(11) Verbal Logical Selection632
(12) Verbal Classification660
(13) Verbal Analogies886
(14) Dot Pattern Perception.....	.436
(15) Overlapping Shapes.....	.413
(16) Non-Verbal Analogies.....	.581
(17) Paper Formboard859
(18) Spatial Symbols476
(19) Non-Verbal Classification762

The *g* loadings obtained by the Thurstone method of analysis differed from the above only in detail.

It seemed, therefore, that Verbal Analogies was the best single test of *g*, its loading being .886, and it was used as the *g* axis in the later analysis.

After taking out this *g* factor, a factor of some size followed in the Mechanical Arithmetic Test. This factor was present also in the Non-Mechanical Arithmetic Test, and yet it was not significantly present in the other tests. This evidence, together with that of reports of other investigators, pointed to this being a number factor. The test of Mechanical Arithmetic was loaded with it to the extent of .459, the *g* loading of the test being .490. This test was adopted for the axis of the *n* or number factor.

¹ SPEARMAN : *The Abilities of Man*, Ap. xvi.

A further factor was found common to all the verbal tests, independently of whether they involved memory, imagery or speed, yet this factor was not significantly present in the non-verbal tests. In the light of previous¹ work this factor was taken to be verbal. The loadings of this factor were high in the Finnish and German word tests. The Finnish test was chosen for the v axis.

The tests of logical selection, classification, and opposites were loaded to the extent of $\cdot 346$, $\cdot 259$, and $\cdot 251$ with another factor, which was not present in the other tests used. In view of previous work, it seemed most probable that this was a factor involving a special logical ability, and it was decided to use the logical reasoning test as an axis for this factor. Unfortunately this test had to be discarded in the final investigation as it entailed a considerable amount of time—(and if the time had been shortened the items involved would have been too few to warrant many conclusions)—the length of testing time being strictly limited in the final groups of children.

A factor common to all the spatial tests, including the Passalong Test, with the exception of Overlapping Shapes, was also detected. The factor was present even in such tests as Non-Verbal Classification, where little scope for spatial imagery was given. It was at its greatest in the Paper Formboard Test, where it was $\cdot 326$, and this test was selected for the axis for this factor in the later work. This factor has been termed the spatial factor or S_1 .

Still another factor was common to the Spatial Imagery, Verbal Imagery, and Paper Folding tests, its loading being the greatest, $\cdot 379$ in the test of Spatial Imagery. It would seem reasonable to interpret this factor as one involving Imagery. The Spatial Imagery test was used as the axis for this factor in the later work, and this factor has been signified by S_2 .

From this analysis, tests which could be used for the g , n , v , S_1 , S_2 , and logical factors were obtained. The factorial make-up of these tests was ascertained and further, it was found that within the variations of age to be tested (i.e., ten to eleven years, and the School Certificate children of fifteen to sixteen years of age) there was no significant correlation with age.

III.—THE MAIN EXPERIMENT.

The subjects.—It was thought desirable to have a variety of types of schools and districts from which the children were drawn to prevent

¹ STEPHENSON : *Journal of Educ. Psych.*, 1931.

such a factor as locality influencing the results. Children from boys', girls', and mixed schools were included, drawn from city, small factory town, and rural environments.

The outside criteria.—In order to have this variety, the results of two examining bodies at the scholarship standard were taken, but one examining body covered all the children of the older group.

The procedure.—So that the chance conditions influencing performance in the tests would be as similar as possible to those operating during the examination, the testing was done as near as possible in point of time to the actual examinations. It was realised that the time of these children was very precious, and the testing time had been reduced as far as was compatible with efficiency in the preliminary work. Finally, permission was obtained to test the younger group in the week before they took their scholarship examination, and the older one immediately after the School Certificate examination.

(i) *The Junior Group.*

District A.—One hundred and twenty of the children described in the above section took the scholarship examination in this district.

Two parallel sets of papers were used by this authority, one for children whose ages fell between 10 years 11 months and 11 years 11 months, and another slightly easier set for those between 9 years 11 months and 10 years 11 months at the time of testing.

The results in the English and Arithmetic examination were available for all of the 120 children tested, but the results of the Intelligence Test were only obtained for twenty-one candidates in the Senior and for forty-one in the Junior Sections.

The analysis of the results of the 10y. 11m. to 11y. 11m. Section.

The correlations between all the nineteen tests given and the three papers in the scholarship examination were worked out by the Rank method if the numbers involved were few (e.g., in the Scholarship Intelligence Paper), and by the Product Moment method in the other cases. These correlations were then subjected to an analysis by the Thurstone method. The rotation for the *g* factor was, in the light of the preliminary work, made through the test of Verbal Analogies, and the *g* loadings of the papers in the examination were found to be :

	<i>g</i>	<i>n</i>	<i>v</i>
Arithmetic	·541	·409	·076
English	·466	·227	·535
Intelligence Test.....	·803	— ·143	·575

From these loadings it can be seen that the relative importance of the three factors was in the order of *g*, *v*, and *n*. Success in the Intelligence Test, which was verbal, depended almost entirely upon *g* and *v* in the ratio of about 8 to 6. In this group of children no allowance for differences in teaching was possible, for although one school provided a considerable number, most of the children had been selected from several different classes and from different schools, and the numbers from each school were too small to be considered separately. As a result of this, a small factor arising from differences in teaching and attitude to work seems to have arisen—thus the teacher who excels in making her children do well in Arithmetic seems likely to make them do so in English also. This additional loading would naturally fall into the first factor involving teaching taken in the analysis, after the *g* factor, which is presumably independent of teaching, had been removed. Accordingly the first factor taken after *g*, namely *n*, is a composite factor including this special number ability and the factor due to differences in teaching.

Success then in the English Paper seems to depend primarily upon the *v* factor (loading ·535) and the *g* factor (loading ·466), and to a smaller extent upon the factor probably arising from differences in teaching (loading ·227).

In the Arithmetic Paper, in which it must be remembered are both mechanical and problem sums, success depends upon *g* (·541), *n* and the teaching factor (·409), and *v* (·076).

Summing up very tentatively in view of the small numbers involved, it would seem that the three most important factors making for success at this stage are *g*, *n*, plus a teaching factor, and *v*. They are roughly in the ratio of 3 : 1 : 2 respectively.

The analysis of the results of the 9y. 11m. to 10y. 11m. Section.

The results for the younger and larger section were obtained by exactly the same method. The loadings for the parallel, although not identical papers, are given on the following page:

	<i>g</i>	<i>n</i>	<i>v</i>
Arithmetic	·434	·550	— ·007
English	·116	·481	·505
Intelligence Test.....	·391	·320	·437

The difficulty noted in the Senior Section arising from the composite character of the second factor obtained, termed *n*, is found here in an intensified form. The difference in attainment and attitude to work assumes considerable importance in this table.¹ Since these younger children have had but five years in school, their standard of attainment in the 'tool' subjects of English and Arithmetic spreads its influence into the admittedly entirely *verbal* Intelligence Paper in the examination and into Mechanical Arithmetic. With this important aspect of the table in mind we may consider the influence of the general and special abilities upon success, as for the other group.

Success in the Intelligence Paper again depends upon the verbal factor to a considerable extent (·437) while the *g* factor is of about half the importance found for the older children, being (·391.) This difference may be accounted for in part by the fact that the Intelligence Paper was different and by the intrusion of the factor arising from differences in school attainment. Success in the Arithmetic Paper similarly depends upon *g* (·434), but to an even greater extent upon the factor termed *n*, which involves 'number' and differences in school teaching and attitude. In the English Paper the verbal factor is again of first importance, closely followed by the school factor, for seemingly the success of these younger people is dependent to a considerable extent upon the teaching they receive, and the *g* factor is of small importance in this paper.

It would seem, then, that the analysis for the younger group is much less complete than for the older group, the factors affecting success seeming to be more complex. Summing up, it may be suggested that the three most important factors are again *g*, *n* plus a teaching factor, and *v*. They are roughly in the ratio 1 : 1·5 : 1.

District B.—In this district three hundred children were tested. They were drawn from sixteen city schools and were between ten and twelve

¹ Evidence of very different standards of teaching in the junior schools was obtained, for although the average scores of the children from different schools were quite similar in the tests which were independent of teaching, the average scores differed to the extent of 10 per cent in Arithmetic and 12 per cent in English in the examination results.

years of age. These children had been selected by an internal examination by the heads of their respective schools for the scholarship examination.

The intercorrelations of the papers given by this authority were analysed as in District A, and the following factor loadings resulted :

	<i>g</i>	<i>n</i>	<i>v</i>	<i>Time allowed.</i>
Mental Arithmetic.....	·69	·03	·02	10 mins.
Mechanical „	·61	·10	·45	15 „
Problem „	·72	·13	·07	45 „
English Composition.....	·48	—·40	·49	30 „
„ Comprehension	·69	—·31	·11	45 „
Intelligence Test	·82	—·47	·20	60 „

The results of each of the above papers were available separately, and since the number of children was large, this group proved to be of great value from the point of view of this research. The factorial interpretation of success could be made with greater confidence than from the results obtained in District A.

That success in this Intelligence Test depended almost entirely upon the *g* factor is abundantly evident, but the *v* factor, although relatively small, is still playing its part. In this analysis the *v* factor was obtained by rotation through English Composition, and the resultant factor seems to have a composite character involving not only the verbal factor, but also a factor which was discussed in dealing with District A, and termed a 'school' factor. Success in Mental Arithmetic seems to depend to a great extent upon *g*, while the *n* and *v*, and probably 'school' factor, play relatively unimportant parts. The abilities operating in Mechanical Arithmetic present a somewhat different balance, for although *g* is still dominant, the 'school' factor and the 'number' factor are both playing a more important role. Proceeding to the Problem Paper, the influence of *g*, as was to be expected, assumes the dominant position while the 'number' factor is second in importance, and that of 'school' is present, but admittedly small.

In English the results are more instructive than those for District A, in that here the results for Composition and Comprehension are considered separately. In Composition the *g* loading is comparatively low, the verbal and 'school' factor together being first in importance. In the Comprehension Paper *g* is again dominant with a loading of ·688, while

the combined verbal and 'school' factor are present, but of relatively little importance.

Summarising for the analysis of the results for this district, it might be claimed (from the Σ of h^2) that about 60 per cent of the causes of success lie in the factors obtained, the analysis being most complete for the Intelligence Paper. Success again depends primarily upon the factors g , v , n , and a 'school' factor, but in this case, because of the length and high g loading of the Intelligence Test and the importance given to Comprehension in the English Paper, and to the Problem Section in Arithmetic, the relative importance of the factors differs from that in District A, the ratio of the factors g , n , and v being 15 : 1 : 5.

From this part of the research one very obvious and important conclusion must be stressed—namely, that the relative importance of general and special abilities upon success in the Scholarship Examination varies enormously with the make-up of the papers set, and further, even between papers which purport to cover the same ground; thus a paper termed English may be principally testing a verbal ability, and secondly a 'school' factor if the compositional aspect holds a dominant position, while the g factor may be of little importance. In another examination the English Paper may stress Comprehension rather than Composition, and the g factor becomes more important than the verbal in such a case. The greater the preponderance of the problematic type of sum in Arithmetic, the greater will be the relative importance of g , while the mechanical aspect of the subject will tend to involve the n and 'school' factor to a greater relative extent. Even in the Intelligence Tests great variation in the g loading will be found, the danger seeming to come from the tendency to involve a high loading of the verbal factor in some of the types of tests used. The verbal factor loadings of some such types of verbal tests were obtained in this research. Thus an 'Opposites' Test had a v loading as high as .587, Verbal Classification of .200, and Logical Selection of .270, and such are frequently included in an Intelligence Test.

As has been shown, success at the beginning of the secondary school career, in the districts studied, depended in the main upon the general factor and two special factors, namely v and n . Other special factors which may play an important part in affecting the extent to which a child will benefit from such education are ignored.

In the factor analysis of the tests used in the preliminary work, as described in the first part of this paper, the presence of a 'spatial' and 'spatial imagery' factor had been detected. The spatial factor had been greatest in the Paper Formboard Test, and the spatial imagery factor

greatest in the Paper Folding Test, and these tests had been used respectively as axes for the respective factors. The spatial tests, with the exception of the individual Passalong Test, had been given to the children whose scholarship papers had been analysed in both Districts A and B, and the results obtained were analysed by the same method as was applied to the scholarship papers. The Verbal Analogies Test, having already been found to be highly loaded with g , was used as the g axis. The Paper Formboard and Paper Folding Tests were used as axes for the spatial and spatial imagery factors respectively, following the findings of the preliminary work. The resultant factor loadings are given below :

	g	S_1	S_2
Analogies Test	·718	—	—
Paper Formboard Test	·234	·720	—
Paper Folding Test	·438	·530	·225
Correlate Education Test	·404	·244	—·308
Field Test	·407	·399	—·088
Spatial Symbols Test	·405	·040	—·344
Tracing Test	—·016	·383	—·232

From the above table it seems that there is a factor S_1 operating in all the spatial tests used. It therefore seems probable that this a spatial factor. The third factor, S_2 , is operating in the Paper Folding Test, and it is suggested, as in the preliminary work, that this may be a factor involving spatial imagery.

Finally, it may be concluded from the results of this part of the investigation that success at the beginning of a secondary school career depends upon g , v , and n , and a 'school' factor, while other factors such as S_1 , which were present in the tests (cf. the last table), hardly function in the Arithmetic, English, and Intelligence Papers of the authorities which were considered in this investigation.

Part II of this article will appear in the next issue.

PERFORMANCES IN THE CONSTITUENT PARTS OF A HIGHER SCHOOL CERTIFICATE EXAMINATION IN CLASSICS.

By JAMES A. PETCH.

- I.—*Introduction : the nature of the data.*
- II.—*Performances in the three subjects, Greek, Latin, and Greek and Roman History.*
- III.—*Performances in the parts of the subjects.*
- IV.—*Performances in the subjects and parts thereof at successive attempts.*
- V.—*Summary.*

I.—INTRODUCTION.

1.—The three subjects Greek, Latin and Ancient History form perhaps the most closely knit of all the subject groups which go to make up the curricula of schools. To what degree is ability in one of these subjects found to go with ability in one or both of the others so far as concerns school pupils?

2.—A public examination gives a wider field of enquiry than does a single school, a field too in which the influence of local or individual peculiarities is reduced. The results for three successive years in Group I (Classics) of the Higher School Certificate Examination of the Northern Universities' Joint Matriculation Board¹ have therefore been analyzed with a view to seeking some more or less quantitative answer to the following questions:

- (1) How far does ability in one classical subject go with ability in another classical subject?
- (2) How far does ability in one portion of a classical subject go with ability in another portion of that same subject or in a portion of another classical subject?
- (3) How far does ability in the subjects as a whole and in separate portions thereof remain constant from one examination to the examination of a subsequent year?

¹ I am indebted to the Board for permission to publish this paper, which is based upon material in its possession, and to Professor C. W. Valentine for valuable suggestions regarding the presentation of it.

Before an answer to these questions can be sought, however, it is necessary first to note certain limitations which condition the enquiry.

3.—Firstly the "ability" in question is ability as shown in an examination by answers to certain papers set on a specific occasion. The extent to which such "one-occasional" ability may coincide with ability shown over a longer or shorter period of preparation is outside the scope of this paper (but see para. 5 (e) below).

4.—Secondly the ability is measured not against any absolute standard, if such there be, but in terms of the order of merit in which the candidates were placed by the examiners; "rank correlations"² are employed throughout.

5.—Thirdly the efficiency of the examination which has produced the results must be considered to see how far consistency of verdict may be expected, how far variations are to be more properly ascribed to examiners than to candidates. In the absence of any absolute standard it is impossible to give a final answer. The following points do, however, help towards an informed opinion.

(a) In the three years reviewed all candidates in Group I offered the same three subjects, namely (i) Greek, (ii) Latin, (iii) Greek and Roman History. In Greek and in Latin there was a choice allowed to candidates only in so far as every candidate had to offer one of two prescribed prose texts and one of two prescribed verse texts, and there was a choice of question in Paper III in both Greek and Latin (General and Literary Questions). There was no choice of period in Greek and Roman History, but there was a choice of question.

(b) The syllabuses in the subjects were not changed during the period under review apart from (i) change of texts prescribed for Greek and for Latin, (ii) rotation of periods prescribed for Greek and Roman History. The first of the three successive years was, however, the first year in which these particular syllabuses and this particular arrangement of papers were in force; even so, the syllabuses then introduced were modified, not completely new, syllabuses.

(c) Throughout the three-year period there were no changes in the allocation of marks as between the various papers and sections of papers (Table I). Moreover it is the constant aim of the examining body to ensure that the papers set in one subject in one year approximate as closely as possible in standard to those set in other subjects in the same year and in the same subject in the previous year.

² Only the greatest "probable error" is given for each table or for each group of coefficients where N is constant or nearly so.

TABLE I.
PAPERS, MARKS AND EXAMINERS.

Section and Marks Allotted.		Examiner responsible.		
		1st year.	2nd year.	3rd year.
GREEK	Paper I (Prose : 28% marks) ..	A	C	C
	(Grammar : 7%) ..	A	C	C
	Paper II (Unseen Translation : 35%)	B	A	D
	Paper III (Prescribed Texts : 16%)	B	A	D
	(General Questions : 14%)	B	A	D
LATIN	Paper I (Prose : 28%) ..	E	E	H
	(Grammar : 7%) ..	E	E	H
	Paper II (Unseen Translation : 35%)	F	F	F
	Paper III (Prescribed Texts : 16%)	G	H	H
	(General Questions : 14%)	G	H	E
GREEK & ROMAN HISTORY	Paper I (Greek History : 50%) ..	J	K	L
	Paper II (Roman History : 50%) ..	K	L	K

(d) The arrangement of the papers in each subject and the distribution of the work of marking the scripts is shown in Table I. There was, therefore, a considerable degree of continuity in the membership of the examining panels and in the apportionment of work. Further there is no question of any variability in standard as between examiner and examiner in any one year; each year all answers to any particular portion of a paper were marked by one examiner. The allocation of individual portions did vary somewhat from year to year but in no year did one examiner mark, say, some of the Latin Proses and another examiner the remainder.

(e) The general "reliability" of any examination is always a matter for discussion. A comparison of the estimates submitted by the schools before the examination with the verdicts passed by the examiners does, however, assist towards a valuation of this factor, with the proviso that the school estimates are based upon the work of the candidates over the whole period of preparation, *i.e.*, at least two years normally, while the examiners' verdicts are, as stated in para. 3, primarily verdicts upon ability shown on a specific occasion and presumably under limiting conditions. The groups of classical candidates entered for this examination by individual schools are generally small and the varying methods and standards employed in arriving at and stating the estimates prevent a comparison with the examination results *en bloc*. In the

smaller groups of three and four candidates perfect agreement as to order of merit as between estimates and results was by no means uncommon. Weighted mean coefficients of correlation between school estimates and examination results for the three subjects for the larger groups (nine to nineteen candidates) over the three years were: Greek 0·87, Latin 0·85, Greek and Roman History 0·61 (P.E. $\pm 0\cdot06$). In Latin and Greek, therefore, the examination results agreed very well with the school estimates of the order of merit for the various groups; in Greek and Roman History they agreed no less than they do for other subjects of the "essay" type.

TABLE II.

CORRELATION BETWEEN PERFORMANCES IN CLASSICAL SUBJECTS.

	<i>Number of Candidates.</i>	<i>Greek - Latin.</i>	<i>Greek- Greek and Roman History.</i>	<i>Latin- Greek and Roman History.</i>
1st year	189	0·90	0·65	0·65
2nd year	158	0·87	0·57 ¹	0·58
3rd year	151	0·91	0·60	0·64

¹ ($\pm 0\cdot04$).

II.—HOW FAR DOES ABILITY IN ONE CLASSICAL SUBJECT GO WITH ABILITY IN ANOTHER CLASSICAL SUBJECT?

6.—Table II gives the coefficients of rank correlation for the three years reviewed when the order of merit in one subject is compared with the order in each of the other two subjects of the group. Some basis of comparison is required. The most popular entry in "modern" subjects is for English, French, and History, in "science" subjects for Pure Mathematics, Applied Mathematics, and Physics. The correlations for a random group of 162 candidates offering these three modern subjects in one year were: English-French 0·49, English-History 0·55, French-History 0·48 ($\pm 0\cdot04$). For a random group of 167 offering these three science subjects the coefficients were: Pure Mathematics-Physics 0·69, Pure Mathematics-Chemistry 0·61, Physics-Chemistry 0·67 ($\pm 0\cdot03$).

7.—The combination French, German, and History, as it comprises two languages and an "essay" subject, may be considered to provide

a group of Arts subjects better for comparative purposes. The coefficients for the sixty-one candidates offering this group in one year were: French-German 0.82, French-History 0.49, German-History 0.44 (± 0.07).

8.—If other subjects offered are ignored, 130 candidates offered French and German in one year; the coefficient of rank correlation was 0.78 (± 0.02). Thirty-four candidates offered French and Spanish; the coefficient was 0.72 (± 0.06).

9.—Thus there is a high correlation between results in Greek and results in Latin, appreciably higher than the correlation between results in any two modern languages. Correlation between results in Greek and Roman History on the one hand and results in the classical languages on the other is not so high. Linguistic ability in Classics is not markedly linked with ability in Ancient History.

III.—HOW FAR DOES ABILITY IN ONE PORTION OF A CLASSICAL SUBJECT GO WITH ABILITY IN ANOTHER PORTION OF THAT SAME SUBJECT OR IN A PORTION OF ANOTHER CLASSICAL SUBJECT?

10.—To what extent does the high correlation between Greek and Latin results hold good? Is it confined to the results in the two subjects as wholes, these results representing a balancing up of series of unequal achievements? Or is a candidate who offers up a good Greek Prose likely to offer up also a good Greek Unseen, a good Latin Prose and so forth?

11.—It is not possible within the limits of this paper to give separately for the three years the individual tables of the coefficients of correlation between the five portions into which each of the two language subjects falls (Table I). In Table III, therefore, are given the weighted means as follows:

(a) in columns 1 to 10 the mean coefficients of correlation between the portion specified on the left and the portions specified at the head of each of these columns;

(b) in column 11 the mean correlation of the portion specified on the left with the nine other portions in columns 1 to 10;

(c) in column 12 the Standard Deviation.

12.—As regards the details of which Table III is a summary:

(a) each year Greek Unseen gives the highest mean correlation with the other four portions of Greek, Greek General Questions the lowest; after Greek Unseen ranks Greek Prescribed Texts, third comes Greek Prose with Greek Grammar close behind;

TABLE III.
CORRELATION BETWEEN PERFORMANCES IN THE SPECIFIED PORTIONS (WEIGHTED MEANS FOR THREE YEARS).

	1	2	3	4	5	6	7	8	9	10	11	12
	GREEK.						LATIN.					
	Prose.	Gram.	Unseen.	P. Texts.	Gen. Q.	Prose.	Gram.	Unseen.	P. Texts.	Gen. Q.	Mean.	S.D.
GREEK	Prose ..	—	0.51	0.71	0.53	0.46	0.51	0.70	0.48	0.40	0.55	0.11
	Grammar ..	0.51	—	0.59	0.56	0.50	0.54	0.56	0.52	0.45	0.53	0.04
	Unseen ..	0.71	0.59	—	0.70	0.55	0.60	0.89	0.62	0.43	0.64	0.12
	P. Texts ..	0.53	0.56	0.70	—	0.57	0.58	0.68	0.76	0.43	0.59	0.10
	Gen. Ques. . .	0.45	0.50	0.55	0.57	—	0.43	0.53	0.53	0.65	0.51	0.07
LATIN	Prose ..	0.68	0.49	0.67	0.50	0.41	0.54	0.71	0.48	0.37 ¹	0.54	0.12
	Grammar ..	0.51	0.54	0.60	0.58	0.43	—	0.60	0.53	0.38	0.52	0.08
	Unseen ..	0.70	0.56	0.89	0.68	0.53	0.60	—	0.62	0.44	0.64	0.12
	P. Texts. . .	0.48	0.52	0.62	0.76	0.53	0.53	0.62	—	0.43	0.55	0.09
	Gen. Ques. . .	0.40	0.45	0.43	0.43	0.65	0.38	0.44	0.43	—	0.44	0.08

¹(±0.03)

(b) Latin Unseen always has the highest mean correlation with the other four portions of Latin, Latin General Questions the lowest, but the relative positions of the remaining three Latin portions are not constant ;

(c) in each year Greek Unseen gives the highest mean correlation with the five Latin portions while Greek Prescribed Texts rank second, Greek Prose third, the remaining two Greek portions showing little differentiation ;

(d) of the Latin portions Latin Unseen gives the highest mean correlation with the Greek portions ; the order of the means for the remaining four Latin portions fluctuates from year to year, save that Latin General Questions always occupies the fifth place ;

(e) when the mean correlation of each portion with the remaining nine portions, Greek and Latin, is considered, either Greek Unseen or Latin Unseen gives the highest mean each year ; each year Greek Prescribed Texts comes third, while Latin Prescribed Texts is fifth, fifth, and sixth ; Greek Prose (fourth, fourth, eighth) fluctuates with Latin Prose (ninth, seventh, fourth) ; each year Latin General Questions is tenth.

TABLE IV.

CORRELATION BETWEEN PERFORMANCES IN PORTIONS: COEFFICIENTS AND WEIGHTED MEANS (ALL THREE YEARS) OF AT LEAST 0.70.

	1st year.	2nd year.	3rd year.	Weighted mean (all three years).
Greek Prose-Greek Unseen	0.76	—	0.73	0.71
Greek Prose-Latin Prose	—	—	0.73	—
Greek Prose-Latin Unseen	—	—	0.73	0.70
Greek Unseen-Greek P. Texts	0.70	—	0.72	0.70
Greek Unseen-Latin Prose	—	—	0.77	—
Greek Unseen-Latin Unseen	0.88	0.88	0.91	0.89
Greek P. Texts-Latin Unseen	—	—	0.70 ¹	—
Greek P. Texts-Latin P. Texts	0.79	—	0.82	0.76
Greek Gen. Ques.-Latin Gen. Ques. ..	0.71	—	—	—
Latin Prose-Latin Grammar	—	—	0.72	—
Latin Prose-Latin Unseen	—	—	0.80	0.71

¹ (± 0.03).

13.—In Table IV are given all instances in which the coefficient of correlation for two portions was at least 0.70. Only one pair appears in all three years—Greek Unseen and Latin Unseen. As, moreover, the

three coefficients for this pair are never less than 0.88, it would appear that in the examination a good (or bad) performance in Greek Unseen is generally accompanied by a good (or bad) performance in Latin Unseen.

14.—Three pairs appear in two years; in each case in the missing year the coefficient was well over 0.60. It would appear, therefore, that there is a significant degree of correlation between the following pairs:

- (2) Greek Prescribed Texts and Latin Prescribed Texts;
- (3) Greek Prose and Greek Unseen;
- (4) Greek Unseen and Greek Prescribed Texts.

To this list may be added the two pairs (5) Latin Prose and Latin Unseen, (6) Greek Prose and Latin Unseen; in each case the means are greater than 0.70 and the terms missing from Table IV approximate closely to the limit adopted.

15.—Negatively: (a) in no year was the correlation between Greek Prose and Latin Prose one of the five highest, the mean being 0.68;

(b) in every year the lowest correlations were those in which either Greek General Questions or Latin General Questions was concerned;

(c) almost uniformly low correlations were given by (i) Latin Grammar and Latin General Questions, (ii) Latin Prose and Latin General Questions.

16.—Before conclusions are drawn from these observations any variables should first be considered. The effect of some possible variables is imponderable—for instance, the standard of the papers (*cf.* para. 5 (c)); one variable of which the influence can at least be estimated is the examiner marking the scripts (Table I). In Table V, therefore, are set out the pairs of portions between which correlation is high or low, together with the examiners responsible for the marking.

TABLE V.
HIGH AND LOW COEFFICIENTS FOR PERFORMANCES IN PORTIONS WITH
EXAMINERS RESPONSIBLE FOR MARKING.

	1st year.		2nd year.		3rd year.	
	Coef.	Exmrs.	Coef.	Exmrs.	Coef.	Exmrs.
(a) Greek Unseen-Latin Unseen ..	0.88	B, F	0.88	A, F	0.91	D, F
(b) Greek P. Texts-Latin P. Texts	0.79	B, G	0.66	A, H	0.82	D, H
(c) Greek Prose-Greek Unseen ..	0.76	A, B	0.64	C, A	0.73	C, D
(d) Greek Unseen-Greek P. Texts	0.70	B, B	0.69	A, A	0.72	D, D
(e) Greek Prose-Latin Unseen ..	0.69	A, F	0.69	C, F	0.73	C, F
(f) Latin Prose-Latin Unseen ..	0.66	E, F	0.68	E, F	0.80	H, F
(g) Greek Prose-Greek Gen. Ques.	0.49	A, B	0.47	C, A	0.37	C, D
(h) Latin Grammar-Latin Gen. Ques.	0.42	E, G	0.33 ¹	E, H	0.39	H, E
(j) Latin Prose-Latin Gen. Ques.	0.33	E, G	0.35	E, H	0.43	H, E

¹ (± 0.05).

17.—Table V, line (a). The high correlation between Greek Unseen and Latin Unseen cannot be entirely ascribed to continuity of examiners. The examiner marking Latin Unseen was unchanged, but three different examiners marked the Greek Unseens. In view of this and the facts contained in para. 13, it may be concluded that there is something in the nature of a "facility for Unseens" which is shown by the same candidate both in Greek and in Latin.

18.—Table V, lines (b), (c), (d). In each of these three lines the second year shows the lowest coefficient and in each line examiner "A" appears in the second year. His marking may, therefore, be taken as influencing the degree of correlation but the extent of his influence seems to vary: in line (b) it appears to be considerable, in line (d) his correlation with himself, as it were, is practically identical with the correlation in the first year, "B" marking both portions, and only slightly less than that in the third year, "D" marking both portions. It may, therefore, be concluded that there is a significant degree of correlation between abilities in these respective portions though in three specific instances the correlation varied in greater or less degree with the examiners who marked the scripts.

19.—Table V, lines (e) and (f). Line (e) calls for little comment: there is little variation from year to year in the coefficients and examiner "F" marked the Latin Unseens each year, his results in the first year corresponding well with those of "A" for Greek Prose, in the other two years with those of "C." Line (f) on the other hand suggests that the high correlation in the third year may be at least partly due to the substitution of "H" for "E." The pair "E, F" are consistent at an appreciably lower level, and the somewhat paradoxical conclusion appears to be that examination results in Latin Unseen correspond as closely to results in Greek Prose as they do to results in Latin Prose (*cf.* Table III).

20.—Table V, lines (g), (h) and (j). Three different examiners marked Greek General Questions, three different examiners marked Latin General Questions; the coefficients are always low. The lowness may, therefore, be attributed to the nature of the work rather than to the standards of the examiners.

21.—As regards the third subject of Group I, Greek and Roman History, when the performances of candidates in Paper I (Greek History) are correlated with their performances in Paper II (Roman History), the resulting coefficients are for the first year 0.58, for the second 0.66, for the third 0.59 (± 0.04); the correlation between the results in the two papers is, therefore, of much the same degree as that between the results in the subject as a whole and the school estimates (para. 5 (e)).

22.—The answer to the question at the head of this section is, therefore, as follows:

(a) A candidate who performs well in Greek Unseen generally does well in Latin Unseen and *vice versa*. If this is due to some "Unseen facility," there is little evidence of any corresponding "Prose facility."¹

(b) The performance of a candidate in Unseens (Greek and Latin) corresponds most closely to his performances in other portions of the subjects, more closely than do his performances in Prose (Greek and Latin). After the Unseens in degree of correspondence come Greek Prescribed Texts. A performance in Greek General Questions or in Latin General Questions bears little relation to what the candidate does in other portions.

(c) The high correlations between the subjects Greek and Latin as a whole (Table II) are not accompanied by equally high correlations between portions of the subjects, Unseen Translation excepted. Generally, therefore, a candidate's performance in Greek or Latin as a whole is the sum of unequal performances in the portions of the subject.

23.—To conclude this section some statistics from a School Certificate Examination of the same Board are given for comparison.

(a) Three hundred and forty-seven candidates offered both Greek and Latin; the coefficient of rank correlation was 0.78 (± 0.01); cf. Table II.

(b) The correlation between performances in Latin, Paper I (Grammar, Unseen, Composition) and in Latin, Paper II (Prescribed Texts) of a random 714 candidates was 0.74 (± 0.01); cf. Table III. For Greek I (Grammar, Unseen, Sentences) and Greek II (Prescribed Texts) the correlation was also 0.74 (± 0.02). In comparing these School Certificate figures with those for the Higher School Certificate the greater sectionalization of the Higher School Certificate results must be borne in mind.

IV.—HOW FAR DOES ABILITY IN THE SUBJECT AS A WHOLE AND IN SEPARATE PORTIONS THEREOF REMAIN CONSTANT FROM ONE EXAMINATION TO THE EXAMINATION OF A SUBSEQUENT YEAR?

24.—Approximately one quarter of all the candidates for the N.U.J.M.B. Higher School Certificate in any one year are making their second attempt at the examination. For Group I the proportion is

¹ It might be suggested that, since the passages set for Unseen Translation are two to three times as long as the passages set for Prose, a candidate may have a better opportunity of doing himself justice in the Unseen Translation paper. As against this, however, it is undoubtedly much harder to ensure that a passage for Unseen Translation shall be really "unseen" for all candidates than to set a passage for Prose from an unknown source.

higher: (i) of the 158 candidates offering Group I in the second year under review, 60 had offered the group in the previous year; (ii) of the 151 offering the group in the third year, 62 had offered it in the second. In considering the successive performances of these candidates the statements in para. 5 are of especial relevance.

25.—In Table VI are given (a) the correlations between the performances in a subject in the first year and performances by the same sixty candidates in the same subject in the second year, (b) correlations for sixty-two attempts in the second and third years.

TABLE VI.
CORRELATIONS BETWEEN PERFORMANCES IN SUCCESSIVE YEARS
(WHOLE SUBJECTS): PARA. 25.

Greek: (a) 1st year-2nd year	0.78
(b) 2nd year-3rd year	0.83
Latin: (a) 1st year-2nd year	0.83
(b) 2nd year-3rd year	0.79
Greek and Roman History: (a) 1st year-2nd year	0.69
(b) 2nd year-3rd year	0.47 ¹

¹ (± 0.07).

26.—Both for Greek and for Latin there is a very fair degree of correlation between performances in successive years though the coefficients are not so high as those of Table II. A factor of unascertained significance may contribute towards this—the amount of time which elapses between a candidate beginning Latin and his beginning Greek, variable from school to school. Unless the two languages are begun together, an unusual occurrence in N.U.J.M.B. schools, at his second attempt the difference between the time the candidate has spent on Latin and the time spent on Greek will, of course, be proportionately less than it was at his first attempt.

27.—As regards Greek and Roman History, Table VI shows a degree of correlation for the first year-second year attempts which is not borne out by the second year-third year attempts, although examiner "J" examined only in the first year together with "K," while in both the second and third years examiners "K" and "L" worked together. It seems safer, therefore, not to stress unduly the higher coefficient.

28.—There remains the second part of the problem with which this section deals—How do candidates fare in successive attempts at

the same portions of the subjects? How do performances in, say, Latin Unseen of one year compare with performances by the same candidates in Latin Unseen of the next year? Table VII gives the correlations for such performances under (a) for the sixty candidates under para. 25 (a), under (b) for the sixty-two candidates under para. 25 (b). After each coefficient are given the examiners concerned.

TABLE VII.

CORRELATION BETWEEN PERFORMANCES IN SUCCESSIVE YEARS (PORTIONS OF SUBJECTS) AND EXAMINERS RESPONSIBLE FOR MARKING: PARA. 28.

	GREEK.				LATIN.			
	(a) 1st year- 2nd year.		(b) 2nd year- 3rd year.		(a) 1st year- 2nd year.		(b) 2nd year- 3rd year.	
	Coef.	Exmrs.	Coef.	Exmrs.	Coef.	Exmrs.	Coef.	Exmrs.
Prose	0.48	A, C	0.67	C, C	0.60	E, E	0.58	E, H
Grammar	0.54	A, C	0.40	C, C	0.42	E, E	0.41	E, H
Unseen	0.85	B, A	0.81	A, D	0.82	F, F	0.84	F, F
Prescribed Texts ..	0.50	B, A	0.59	A, D	0.40	G, H	0.48	H, H
General Questions	0.39	B, A	0.57	A, D	0.52	G, H	0.37	H, E

	GREEK AND ROMAN HISTORY.			
	(a) 1st year- 2nd year.		(b) 2nd year- 3rd year.	
	Coef.	Exmrs.	Coef.	Exmrs.
Paper I	0.57	J, K	0.42	K, L
Paper II	0.52	K, L	0.28 ¹	L, K

¹ (± 0.08).

29.—In para. 17 it was suggested that there was something in the nature of a facility for Unseens holding good both in Greek and in Latin. Further evidence for this is given by Table VII, which suggests that not only is a candidate who is good (or bad) at Latin Unseen generally good (or bad) at Greek Unseen, but that a candidate continues to achieve approximately the same goodness (or badness) from year to year. The uniformly high coefficients in this table for successive attempts at Greek Unseen and at Latin Unseen cannot certainly be ascribed only to the

examiners themselves. "B" marked Greek Unseens the first year, "A" the second; their verdicts upon sixty candidates in successive years produce the high coefficient of 0.85. "F" marked Latin Unseens in all the three years and the two coefficients resulting are 0.82 and 0.84. This Unseen facility is, therefore, constant within narrow limits.

30.—As in para. 22 (a), so here it may be said that there is no evidence of a corresponding Prose facility. Examiner "C" produces, under (b), a much higher correlation between his successive verdicts upon Greek Proses than do examiners "A" and "C," under (a); on the other hand in Latin Prose, under (a), examiner "E" agrees with himself in successive years only as well as he does with examiner "H," under (b).

31.—In the period reviewed seventeen candidates made three successive attempts at the group. The results of an analysis of their performances cannot be stressed; they do, however, bear out generally the conclusions of this section.

32.—The answer to the question at the head of this section is, therefore, that, while there is fair constancy in Greek and Latin performances as wholes, there is no marked constancy as regards the various portions except in Unseen, where the degree of constancy is considerable. In all other portions ability varies much from year to year.

V.—SUMMARY.

33.—Bearing in mind the limitations of the enquiry (paras. 3 and 4) and assuming that there is evidence that the demands of and evaluation by the examiners do not fluctuate greatly throughout the examination in one year or from year to year, the following conclusions emerge.

(a) Ability in Greek is closely linked with ability in Latin, whereas ability in the two languages bears no marked relation to ability in Greek and Roman History.

(b) Equally meritorious performances in Greek and Latin as whole subjects may be resolved into very unequal performances in the various portions of the subjects, except that ability in Greek Unseen is very closely linked with ability in Latin Unseen.

(c) Generally performances in Unseen Translation (Greek and Latin) correspond the most closely to performances in other portions of the subjects, while ability at Unseen Translation is the most stable (i) within the examination of a single year, (ii) from one year to the next, being much more stable than any ability for Prose.

(d) Ability in General and Literary Questions is consonant with very unequal performance in other portions of the language subjects.

34.—For a full discussion of ability in Classics in a wider and more general sense than that used in this paper (para. 3) many more factors

must, of course, be considered. It would, therefore, be unwise merely on the evidence which may have come to light from this enquiry to lay too much stress on what that evidence suggests, namely that in any examination in Classics in which Prose Composition were allowed to play the predominating part the examiners would be stressing an ability which is variable as contrasted with the much greater constancy of ability in Unseen Translation.

RÉSUMÉ.

DES RÉSULTATS DANS LES PARTIES CONSTITUANTES D'UN EXAMEN DU "HIGHER SCHOOL CERTIFICATE" DANS LES LANGUES ANCIENNES.

L'on analysa les résultats pendant trois ans dans les trois branches, le grec, le latin et l'histoire ancienne, et dans les sections de ces branches, dans l'examen du "Higher School Certificate du Northern Universities Joint Matriculation Board." Le coefficient de la corrélation le plus bas entre les résultats en grec et en latin était 0.87. Les résultats ne correspondaient pas de très près aux résultats dans l'histoire ancienne. Le coefficient de la corrélation le plus bas entre la version grecque et la version latine était 0.88; la corrélation entre les résultats dans les autres sections n'était pas si élevée. Environ deux cinquièmes du nombre des candidats se présentèrent à l'examen au moins deux fois. Le coefficient de corrélation le plus bas entre deux tentatives successives en version était 0.81; dans les autres sections des branches la coefficient n'était nulle part au-dessus de 0.67. Il semble y avoir quelque temoignage d'une aptitude spéciale à la version, qui est valable pour les deux langues, et qui reste stable d'année en année.

ZUSAMMENFASSUNG.

LEISTUNGEN IN DEN BESTANDTEILEN EINER PRÜFUNG IN KLASSISCHEN SPRACHEN UND ALTER GESCHICHTE IM HIGHER SCHOOL CERTIFICATE (ENGLISCHEN SCHULABGANGSZEUGNIS).

Man hat Leistungen in drei Fächern Griechisch, Latein, und alter Geschichte und in den drei Teilen dieser Gegenstände in der H.-S.-C.-Prüfung der N.U.J.M.B.-Behörde über einen Zeitraum von drei Jahren analysiert. Der niedrigste Korrelationskoeffizient zwischen Leistungen in Griechisch und Latein ergab sich als 0,87; Leistungen in diesen Sprachen standen nicht in sehr naher Verbindung mit Leistungen in alter Geschichte. Der niedrigste Korrelationskoeffizient zwischen Leistungen in nichtvorbereiteter Übersetzung aus dem Griechischen und nichtvorbereiteter Übersetzung aus dem Lateinischen war 0,88; die Korrelation zwischen Leistungen in anderen Teilen war nicht so hoch. Annähernd zwei Fünftel der in Frage kommenden Kandidaten machten mindestens zwei Versuche beim Examen. Der niedrigste Korrelationskoeffizient zwischen aufeinander folgenden Versuchen im nichtvorbereiteten Übersetzen war 0,81; für keinen anderen Teil der Fächer war der Koeffizient höher als 0,67. Es gibt anscheinend Beweise von besonderer Fähigkeit für nichtvorbereitete Übersetzung, die in anderen Sprachen vorhanden ist und von Jahr zu Jahr bleibt.

THE FACTORIAL ANALYSIS OF HUMAN ABILITY.

By GODFREY THOMSON, D.Sc., Ph.D., Professor of Education in the University of Edinburgh. (University of London Press, Ltd., 1939, pp. xvi+326. 16s.)

THE theory of factors made its first appearance as an offshoot of the mental testing movement ; but it has come to play much the same part in modern psychology as the theory of linear operators in modern physics. Indeed, as branches of mathematics both are closely akin. Prof. Thomson was one of the first to introduce standardized tests of intelligence and attainments into the regular examinations for junior county scholarships and the like ; and, in the endeavour to put such examinations on a more scientific basis, he was led at the very start to scrutinize the nature and the defects of the procedure from a mathematical standpoint.

What are the factors that chiefly influence an examinee's performance in such tests ? How can they be discovered, described, and assessed ? To answer such questions numerous methods of ' factor-analysis ' have from time to time been put forward ; and a further problem therefore arises : which out of all these methods is the best ? As Prof. Thomson observes, " before any practical use can be made of such analysis, a thoroughgoing examination of its mathematical foundations is necessary." Accordingly, he has himself undertaken a comprehensive review of the various basic principles proposed, and carefully compared the merits and the defects of each procedure.

The results of his work fully justify the eagerness with which they have been awaited. To the non-mathematical factor-analysis wears a formidable aspect. But Prof. Thomson's book can be understood by students with little or no mathematical knowledge beyond what is ordinarily imparted in a secondary school. At the same time even professional statisticians will welcome, not merely the succinct compendium of formulae and proofs brought together in the mathematical appendix, but also the new light thrown on old problems and on familiar solutions at almost every turn of the page. Throughout the author reveals himself, not only as a brilliant mathematician, but also as a brilliant teacher. The structure of the book, the compact and lucid style, the ingenious devices used to illustrate geometrical conceptions and arithmetical proofs, make the whole work easy to follow and brimful of interest.

The introductory explanations are simplified by assuming that sampling errors may be provisionally ignored. On this basis Part I deals with ' the analysis of the tests.' As his starting point, Prof. Thomson takes the simple hierarchy of test-correlations on which the two-factor

theory was founded. To the pioneer work of Prof. Spearman he pays a generous tribute. "The main idea which still, rightly or wrongly, dominates factorial analysis was enunciated by him, and practically all that has been done since has been either inspired or provoked by his writings." He then describes how, at a very early period, it became necessary to contemplate group factors for more specialized abilities, in addition to the single central factor popularly termed intelligence, and how the subsequent investigations of Maxwell Garnett and others led to the idea of a multiple-factor analysis as a natural extension of the initial two-factor procedure. The more elaborate methods of Thurstone, Hotelling, and Holzinger are then examined in a way that is at once clear and impartial.

Part II deals with the estimation of factors by regression and other methods. Here and elsewhere Prof. Thomson makes a fruitful use of some of the neat expedients devised by his colleague, Dr. Aitken, for solving what are essentially problems in matrix algebra. Once again the technical explanations are most skilfully treated. The essential principles are introduced as they are needed; and, before he reaches the mathematical appendix, the student will have realized how a novel branch of algebra, unexpectedly easy in its more elementary applications, simplifies and abridges both exposition and investigation.

Part III discusses the complications introduced by sampling and selection, and considers their effects on two-factor and multiple-factor analysis. In Part IV we reach a topic of special interest to the educationist—the proposal to correlate persons instead of tests. As Prof. Thomson points out, the device of correlating persons, with or without the additional idea of extracting factors and determining types, has long been in use; but it is only during the last two or three years that it has come to the front as a topic of controversy. He first reviews experiments in which the general factor for persons forms the chief centre of interest. Children are asked to rank pictures, literary extracts, school subjects and the like according to their preference; examiners are asked to rank essays or examination papers according to their intrinsic merit. How far are the judgments consistent, and what are the causes of the agreement and the disagreement? A number of experiments of this kind have been carried out at the London Day Training College, at University College, and elsewhere, and reported in this *Journal*. In view of their recent applications to so many practical problems in both educational and vocational psychology, Thomson's critical examination of such methods, and of the incidental difficulties to which they give rise, is a welcome contribution.

A chapter is devoted to the relation between test-factors and person-factors; and in his final chapter, headed 'stop press,' he returns to the issue again. The suggestion that, whether we start by correlating persons or by correlating tests and traits, the secondary factors and the corresponding types will nevertheless remain approximately the same, he accepts only with definite reservations; and he points out, with perfect justice, that any endeavour to generalize an equivalence obtained under special conditions and by special methods must raise a good many supplementary problems that have yet to be solved.

To the educationist, however, the most interesting section of all will be the last. Part V brings the volume to a close with a full discussion of 'the interpretation of factors.' Spearman's *g* and its identification with 'mental energy', Thurstone's 'simple structure' which abolishes *g*, are considered in turn; and a much-needed caution is registered against the common tendency to 'reify factors'. And then, in what (apart from 'stop press') is the concluding chapter, Thomson fills in the earlier sketch of his own 'sampling theory'.

The sampling theory is the most far-reaching of Prof. Thomson's many original contributions to the subject. Briefly, he holds that the effect of any mental test is best described by supposing that each "calls upon a sample of the numerous bonds which the mind can form", i.e., it excites a definite number, but a random selection, of the examinee's "neurone arcs". As every reader will remember, the hypothesis was first put forward as an alternative to Prof. Spearman's theory of two factors. The phenomenon which both are concerned to explain is the hierarchical arrangement that all the early correlation tables seemed to exhibit. According to Thomson this singular tendency is "the result, not of any psychological laws, but simply of the laws of chance."

The conclusions to which this standpoint commits us are of such importance both practically and theoretically that they require examination at some length. Their practical importance will be obvious to every educationist. A teacher examining his class in geometry or French cannot possibly ask his pupils to demonstrate all the propositions they are supposed to know or to translate all the words or sentences they are supposed to understand: he can only ask for specimen replies, and judge from those. Psychologists continually remind him that his class forms only a sample of the total school population, and probably a selected sample at that; but this kind of sampling does not trouble him. What concerns him is the fact that each pupil's answers yield only a sample of his total knowledge. A different test-paper might lead to very different revelations.

Thomson's problem, however, is not so much the sampling of the tests, or even of the contents of the mind, but rather the sampling of its capacities. The theoretical significance of his conclusions is therefore greater still. No one, I fancy, would now question his mathematical deduction. The criticisms most commonly urged are directed rather against the tacit assumptions: can we really accept the implicit principle that the essential working of the brain can be reduced to these simple algebraic formulae? Thomson's reply is briefly as follows. In psychology, as in every other science, we must take the smallest possible number of the simplest possible postulates that can be extracted from facts at present known to us, and see how far we can deduce from them the rest of the facts—including those that have never been suspected and have still to be verified by experiment.

Six postulates seem implied in Thomson's theory. Let me try to formulate them in Thomson's own terms—though perhaps he may protest that my summary gives his picture too definite and too concrete a colouring. So far as they go, I believe every psychologist—including Spearman himself—would accept them. (i) The human brain is a comparatively undifferentiated structure of cellular elements: these elements are so numerous and so similar that for statistical purposes their behaviour may, to a first approximation, be treated as that of a homogeneous population almost infinitely large. (ii) By means of association paths or 'bonds' each cell is connected, directly or indirectly, with every other: no doubt, the lower levels of the central nervous system are loosely organized into so-called 'centres'—forming 'sub-pools' within the total 'pool'; but on the higher levels localization is ill-defined; and, as compared with the members and organs of the body, "the mind shows hardly any structure". (iii) Owing to these interconnecting paths, the energy supplied to, or liberated within, the brain can in theory spread from any one part to any other, so that the total amount of nervous energy can be regarded as forming a single fund. (iv) The stimuli that make up any cognitive test in ordinary use are so numerous that they call into play a mass action of the brain as a whole rather than one or two localized areas. (v) Even so, however, we may distinguish different degrees of complexity both in brains and in tests. Tests differ widely in 'richness', i.e., in the number of 'neurone arcs' they tend to excite, and consequently in the extent to which they are likely to overlap. (vi) Similarly, persons differ widely in the 'richness' of their brains, i.e., in the number of neurone arcs they present to the tests; and, except in rare pathological cases, these differences are general rather than local, characteristic of the tissue as a whole rather than of different areas or layers.

If this is the apparatus, how does it function? How precisely does the initial disturbance, produced by the test-stimuli, distribute itself throughout the brain, before it discharges in muscular action? My own hypothesis—one which would seem largely to reconcile Thomson's views with Spearman's—would begin by laying down four further conditions. The choice of the outgoing path must be determined (i) by the initial arrangement of the various 'neurone arcs'; (ii) by the point at which the incoming currents enter the system; (iii) by the differential resistances met with at successive cell-junctions; and (iv) by the tendency of the currents themselves to establish some kind of equilibrium. Without some such further assumptions we cannot, so it seems to me, offer any clear account of the actual responses to our tests. The scheme is really McDougall's (though, I admit, he himself would be a little shocked at the use I have made of it); and the outcome can be expressed in his own familiar phrase—'an ever-changing tri-dimensional neuergic pattern'¹. If we add (v) a small factor of 'indeterminacy', we can satisfy all that the sampling theory and the hormic theory have, in my view, a right to claim.

And now, with the assumption of less than a dozen constants, we could deduce, as I have elsewhere ventured to suggest, a whole theoretical science of 'neurodynamics', seeking to explain all the measurable phenomena that enter into our tests—the speed of simple and compound reaction-times, the conditioning of reflexes, the eduction of relations and correlates, and even, with matrices to describe the patterns, the perception of Gestalten. A mathematical neurology of this kind may sound far-fetched: but twenty years ago it would have seemed equally far-fetched to explain the efficient or inefficient reception of an orchestral concert by the wiring plan of a receiving-set and a string of differential and integral equations. And the value of all such simplified systems lies, not merely in what they explain, but still more in what they fail to explain.

The hypotheses of Spearman and Thomson take short-cuts towards a simpler solution still. Spearman assumes that the distribution of nervous energy is analogous to the distribution of a limited supply of current from a central generator to various specific engines, "much as the electricity that comes to our houses is distributed to a light-bulb,

¹ "Physiological Factors of the Attention-Process," *Mind*, XII (1903), p. 301. McDougall's 'teleological principle of conation' is really covered by my condition (iv), which is identical with the interpretation of it given by Stout and Avenarius: 'the teleological aspect of conscious activity has its physiological counterpart in the tendency of neural systems to return towards a state of equilibrium' (cf. *Philosophie als Denken gemäss dem Prinzip des kleinsten Kraftmasses*, 1876, p. 45).

a heater, a cooking stove, a radio set and perhaps an electric razor." Thomson prefers to make no specific postulates about directing agencies whatever. All we know about the tests is that each excites a certain specifiable proportion of nerve-elements out of the grand total; which elements they excite, and the consequent patterns produced, must therefore be treated as the result of sheer chance. And so, *for a first approximation*, he assumes—as every statistician would in dealing with “a complex but unknown background of causes”—that the distribution of nerve stimulations obeys the laws of probability and nothing more—like the distribution of the streams of shot poured out in Galton's well-known binomial machine. The immediate inferences are obvious and illuminating: the overlapping of the tests explains their correlations: the saturation coefficients for the tests measure their ‘richness’ or complexity: the factor-measurements for the persons measure the ‘richness’ of the persons’ brains. And the statistical nature of mental testing may be exemplified by tossing dice, and adding up the scores.

From the two antagonistic assumptions made by Spearman and Thomson respectively there follow a series of opposed corollaries which, in spite of their theoretical origin, have an important bearing on educational practice. (i) According to Spearman's interpretation, an individual's g measures the limited amount of mental energy he has at his disposal; according to Thomson, it indicates the limited number of neurone arcs that his brain can call upon—a view that comes very near to Thorndike's ‘quantity hypothesis’ (“quality of intellect depends upon quantity of neural connections”¹). Individual differences in mental energy are not a cognitive but a conative phenomenon—they represent differences in strength of motive-power rather than in strength of intellect. (ii) With each of our tests the correlations are far from perfect, and with certain tests (e.g., tests of simple sensation or movement) the correlations are decidedly low. According to Spearman, this is because each test depends on a specific ability or ‘engine’ peculiar to itself—an ability which (in the case of the simpler processes) far outweighs the influence of the general energy. According to Thomson, an imperfect correlation merely means that not all the pool of neurones has been sampled: for if two tests are of narrow scope they have little chance to overlap. Thus, the two factor theory regards g as a part of the test; and the sampling theory considers the test as a part of g . The practical inferences are plain. If these specific abilities really exist, their measurement should obviously play an important part in educational and vocational guidance; but if not, our practice should be radically altered. (iii) Finally, there is a more

¹ E. L. THORNDIKE: *The Measurement of Intelligence* (1925), p. 415.

abstract corollary. Both Spearman and Thomson admit that the hierarchical tendency is never exact. Spearman attributes the inexactitude to the fact that only a random part of the total pool of *persons* has been sampled; Thomson attributes it to the fact that only a random part of the total pool of *neurones* has been sampled. This at once suggests two crucial experiments that might serve to decide between the two theories.

It will thus be seen that the sampling theory, as developed by Thomson, is primarily opposed to the popular analysis of the mind into a few independent 'unitary abilities', 'factors', or 'traits', in much the same way as the associationist doctrine, as developed by Thorndike, was opposed to the traditional analysis of the mind into a composite of a dozen or so 'faculties'; and so far as it goes, it has much the same arguments in its favour. The reader, however, may naturally ask whether the objections so often urged against Thorndike's associationist doctrine cannot also be urged against Thomson's sampling theory. Can we ever reduce what Stout would call 'noetic synthesis' and what Spearman would call the 'eduction of relations and correlates' to the blind play of mechanical associations? And is it not precisely in the testing of intelligence that the former supersede the latter? "We shall never be able," says Koffka, "to deduce organization or orderly arrangement from the mere indifferent 'spreading' or interpenetration of nerve-currents into innumerable isolated elements." By omitting all reference to the 'integrative action of the nervous system,' has not our physiological picture suppressed something which rules *random* sampling almost entirely out of the question?

One final objection may be raised, which to the educationist might at first seem the most serious of all. 'Dice,' as Prof. Thomson admits, 'cannot be educated.' Yet educability is one of the most striking symptoms of intelligence; and, as he himself points out, the diminution of the hierarchical tendency in correlations obtained from older children seems to mean that 'education imposes a structure on the mind which is absent in the young.' With what may be called the 'directed nerve-current' hypothesis, described above (as distinct from a simple 'nerve-excitation' hypothesis—for physiologically that is what sampling reduces to) it would be by no means difficult to construct a plausible neurodynamic explanation: as every wireless engineer discovers to his cost, certain parts of his electrical circuits display a residual as well as a permanent 'retentivity'; and the phenomena of what is termed (in language apparently borrowed from St. Paul) 'hysteresis'—the varying ease with which iron, steel, and other metals can be magnetized, and the diminishing

succession of magnetic cycles needed for demagnetization—afford very suggestive examples of how quantitative formulae can be derived for what are essentially processes of 'conditioning' and 'deconditioning.' Some such additional principle must surely be included in our scheme, unless we are to ignore the most characteristic effect of mental stimulation. A theory which interprets *g* *solely* in terms of sampling would find it hard to explain the inner working of any test of educability, and harder still to show how the structureless pool of the mind, by merely repeating the sampling process, can ultimately be turned into an organized system of subpools.

But Prof. Thomson does not propose to explain test-performances solely in terms of sampling; and he is by no means unaware of these further problems. Indeed, if, on a first persual, the reader thinks he has discovered a grave objection to the theory, he may be warned to re-read the relevant pages with care: more often than not he will find Prof. Thomson has anticipated the difficulty and even tersely hinted how it can be met. In an introductory work, the primary object of which is to give a comprehensive and impartial survey of the whole subject, suitable for the elementary student, the author has rightly kept his own interpretations modestly in their place. We shall all, however, look forward with the greatest interest to the further elaboration of his views which he seems to promise. Meanwhile, the present volume not only gives a lucid and unbiassed account of the different theories and methods that have hitherto been advanced, but also, as I have tried to show, raises a number of interesting and important issues which factor-analysis has still to answer and on which the thoughtful student can reflect and perhaps research. Both to investigators in this special field and to every teacher or educationist who wishes to think with precision about such questions, the book is indispensable. It is a splendid culmination to twenty-five years' work, and at the same time supplies a solid foundation for work that is yet to come.

CYRIL BURT.

THE SPENS REPORT.

(Report of the Consultative Committee on Secondary Education with special reference to Grammar Schools and Technical High Schools. His Majesty's Stationery Office. 3s. 6d.)

THE latest report of the Consultative Committee is likely to become as famous as the Hadow Report, to which indeed it may be regarded as a sequel. Like its predecessor it has appropriated the name of the Chairman of the Committee, and is already known to the public as the Spens Report. Its terms of reference are narrowly defined. They limit the deliberations to "the organisation and inter-relation of schools, other than those administered under the Elementary Code, which provide education for pupils beyond the age of 11+; regard being had in particular to the framework and content of the education of pupils who do not remain at school beyond the age of about sixteen." The Committee was, therefore, supposed to leave out of account about 85 per cent of post-primary pupils and confine its attention to those who attend secondary schools and vocational schools. As a matter of fact it did not so limit its purview, though it cannot be claimed that it has dealt adequately with the residuary 85 per cent.

Its own special theme, however, together with all matters which throw light upon that theme, is treated with a thoroughness which compels one's admiration. The Report is erudite and at the same time infused with a generous and democratic spirit. The first chapter is a long and learned historical sketch. It traces the development of the secondary school curriculum in England and Wales, and shows that there has been no single and simple line of evolution but rather a series of sporadic attempts to relate school studies to the life of the times. Since the days of the Renaissance the non-local Grammar Schools, as the Report designates our public schools, have exercised a conservative and stabilizing influence over the trend of secondary education in general, and have resisted all efforts to adjust the school curriculum to meet the needs of an ever-changing social environment. They clung stubbornly to the theory that the study of language is the main standby of a liberal education (a theory which is probably sound), and that the study of language is always best approached through Latin and Greek (a theory which is certainly unsound). Other schools, opened by Nonconformist bodies or by private enterprise, afforded a broader and more progressive training, but were regarded as experimental. In any case they lacked

the prestige of the public schools. The upshot is that even at the present day the curriculum of the Grammar Schools, both local and non-local, has been unduly influenced by the traditional curriculum of the public schools.

This historical chapter, for which the Secretary, Dr. R. F. Young, is mainly responsible, is supplemented by a very illuminating appendix for which he is wholly responsible. It deals with the conception of a liberal education. The term "liberal education" derives from the days of Greek and Roman civilization, which was based on slavery. A liberal education was simply the education of a freeman. It excluded the illiberal arts, the mechanical occupations pursued by slaves. As freemen alone were concerned with the government of the state, they alone needed to understand the duties and privileges of a citizen. In these days of universal suffrage it is reasonable that a liberal education should be defined as in the Report: "Such an education as will enable men and women to understand the world in which they live and to contribute to the understanding of its problems." It is interesting to note that two quite different (and quite erroneous) interpretations have been given to the term. One connects it, not with the adjective *liber*, free, but with the noun *libër*, a book, thus making a liberal education mean a bookish education. The other interpretation gives "liberal" the modern meaning of generous, abundant, and makes the term "liberal education" signify one which is broad and comprehensive.

The Committee regard the distinction between liberal and vocational education as invalid. Their Secretary quotes with approval Professor A. N. Whitehead's pronouncement: "The antithesis between a technical and a liberal education is fallacious. There can be no adequate technical education which is not liberal, and no liberal education which is not technical . . . In simpler language education should turn out the pupil with something he knows well and something he can do well." The curriculum of the Grammar School is vocational in that it trains the pupil for a university career, and ultimately for one of the learned professions; the curriculum of the Technical School is liberal inasmuch as it helps the pupil to understand the world in which he lives and in which he will earn a living. One is less surprised therefore to find, among the Committee's recommendations for Grammar Schools, one which is enough to make Arnold and Thring turn in their graves: "In areas in which many pupils leave at, or shortly before, the age of sixteen to obtain clerical posts in industry and commerce, opportunities to acquire the mechanical skills of typewriting and shorthand should be given them after the age of fifteen."

To balance this we find a recommendation that in all secondary schools, whether academic or technical, the training between the ages of eleven and thirteen should be of a broad liberal character.

The Report allots space to a discussion of the influence of Modern Psychology on traditional ideas about the curriculum. It specially mentions the changes in our conceptions of the general effect of special training, and of the nature and growth of intelligence ; and it appends two admirable memoranda which should be of particular interest to the readers of this journal. One, by Professor Cyril Burt, deals with the history of faculty psychology ; the other, by Professor H. R. Hamley, deals with the cognitive aspects of transfer of training. In the light of these memoranda it is clearly seen why Latin lost its dominance over the curriculum and why the curriculum has been enriched by the inclusion of a larger variety of subjects.

The Hadow Report presented us with an ideal of secondary education for all—a very different thing from grammar school education for all. That ideal is reaffirmed by the Spens Report. All post-primary pupils, without distinction of intellectual or social standing, are to receive the secondary education which is their due. *Parity* is one of the key-words of the Report. There is to be parity of status for the schools, parity of opportunities for the pupils, and, with certain reservations, parity of payment for the staffs. All post-primary schools are to be administered under the same Code of Regulations and are to have equal facilities and equal amenities. The classes are to be equally small (thirty is the number mentioned), the premises equally commodious, and the equipment equally generous. No differences are to be allowed unless they are necessitated by differences in the curriculum.

The great practical problem in the education of adolescents will be to take the multitudinous eleven-year-olds (there are close on half a million of them) who emerge every year from the primary schools and distribute them among the various types of secondary schools. It is a question of selection and distribution—if the schools are there. But they are not there ; not as envisaged by the Committee's two idealistic Reports—that is unless we are to believe that the Committee is content with the three categories repeatedly referred to in the Spens Report : the Grammar Schools, the Technical Schools, and the Modern Schools. It suggests changes in the curriculum of Grammar Schools and in the organisation of Technical Schools, but it seems to be content with the proportion of children who attend these three types of school. At present 13·7 per cent of the eligible age-group in the elementary schools pass on to Grammar Schools. The Report suggests raising the average to fifteen—

a trifling change. It hazards no estimate of the number that should go to the Technical High School, the new type of technical school which it is proposed to establish. The numbers attending all schools which may broadly be termed technical are not large. Even in London, which is reputed to be best supplied with these schools, the percentage is only about 4, and that percentage is reached not at 11+ but at 14+. At 11+ the children are distributed among two types of schools only, the Grammar Schools and the Modern Schools, and in the proportion of 15 to 85. Eight-five per cent of adolescents from eleven to fourteen years of age attend Modern Schools, and these schools are frequently referred to as though they were all of the same sort, and were all as homogeneous as the Grammar Schools and the Technical Schools. Are we to take this 85 per cent as constituting the undifferentiated residue, a group awaiting finer classification as soon as we can achieve it; or are we to take it as constituting a final type of school from which there is to be no further split? As a matter of fact in a large number of areas the schools have settled the problem almost for themselves; at any rate with but little help from the central authority. In London and Manchester and in many other areas, rural as well as urban, Modern Schools are of two types, selective and non-selective. And between these two there is a world of difference. They are generally known as Selective Central Schools and Senior Schools respectively. And it is a serious flaw in the Spens Report that it does not properly distinguish between them, and does not give the Selective Central Schools the full credit of their achievements. It brackets both under the common name Modern School. Why they are called "modern" it is difficult to say, for the Selective Central School is nearly thirty years old, and the Senior School is about as old as the elementary school itself. Besides, the common name disguises the wide disparity in their characters. The selective Modern School (to use the nomenclature of the Report) is often, in the intelligence of its pupils, the quality of its work, and the qualifications of the staff, indistinguishable from a good Grammar School; while the non-selective Modern School, having been twice robbed of its brighter pupils, is inferior in its intellectual resources to the top half of the old elementary school.

The inconvenience of using the same term for two such different schools is specially evident in the section dealing with "the combination of a Grammar School with a Modern School in one set of buildings," a policy that is recommended for areas with a sparse population. It is conceded towards the end of the section that "the organisation of a combined school will prove to be more efficient in actual operation if

the modern side is of a selective character." Of course it will, for the school will still be a fairly homogeneous Grammar School; the other alternative produces a multilateral school where the sides are so small that it will be, in every sense, expensive to work.

The detailed recommendations of the Report, since they are by no means likely to be carried out in the near future, need no further consideration here. The Report itself, however, with its ample discussion of general principles, and its fair-minded presentation of *pros* and *cons*, stands as an up-to-date treatise on education, a treatise both informative and inspiring. Though most of it is nominally concerned with a limited group of children—those between eleven and sixteen years of age—the principles discussed and promulgated cover a much wider field. It well expresses certain ideals cherished by the most enlightened and most earnest members of the teaching profession to-day. I specially refer to a unified profession, a unified system of education, and the provision for each child of an education in accordance, not with his capacity to pay for it, but with his capacity to benefit by it. The book does not perhaps always indicate the best way in which these ideals may be embodied in practice, but there will never be in the reader's mind any doubt as to the objective the Committee had in view, nor any doubt as to the rightness of that objective.

P. B. BALLARD.

OUTLINES OF RESEARCHES REPORTED IN THESES PRESENTED FOR HIGHER DEGREES OR DIPLOMAS.

THESE OUTLINES MUST BE SUBMITTED THROUGH THE HEAD OF
THE DEPARTMENT IN WHICH THE RESEARCH WAS CARRIED OUT.

The Sense of Humour in Childhood and Adolescence.

*An Abstract of a Thesis presented for the degree of Ed.B. in the University
of Aberdeen, June, 1938.*

By ALEXANDER LAING, M.A., Ed.B.

A STUDY was made of the differences in the sense of humour (discernment of the laughable) of children from the beginning of the primary-school stage to adolescence. Data were obtained from 709 boys and girls, divided into three age-groups: 7-10, 200; 11-13, 283; 14-18, 226. The subjects causing laughter in the youngest group were investigated orally. The children in the second group wrote for an hour on any of the following subjects: (1) Amusing things I see about me; (2) The most amusing experience I ever had; (3) Comical people I have seen (in the street, in films, etc.); (4) Comical people I have read about. Similar essay-subjects were given to the children in the third and oldest group, together with two new topics, "My sense of humour" and "What makes me laugh," which were included to give the adolescent tendency to reflection an opportunity to express itself, and were actually chosen by 40 per cent of the group. To supplement these inquiries into visual humour, the children's appreciation of verbal wit was studied by asking them to give their best jokes.

Development of the sense of humour was found to run parallel with general intellectual and emotional development. Deviations from the normal and conventional were outstanding causes of laughter in the youngest group. Incidents involving someone's discomfiture (which were also popular in the youngest group) were those most frequently referred to by the children in the second group. The humour of both these groups was predominantly visual; verbal wit was little appreciated. Distinguishing features of the adolescents in the third group were: (a) the marked individual differences in their sense of humour—although many no longer tolerated slapstick farce and preferred witty and realistic comedy, others enjoyed both, and others, again, gave as their best jokes fatuities that were popular among children in the youngest group; similarly, although several relegated the laugh of Hobbes' "sudden glory" to an inferior category, many gave first place to humour of this type; (b) their much greater appreciation of verbal wit; (c) the tendency of the majority to reflect on why they laugh, and of a minority (especially among the girls) to deprecate laughter that is unfeeling.

BOOK REVIEWS.

The Development of Children's Concepts of Causal Relations : By JEAN M. DEUTSCHE. (pp. 104. 9s.)

The Development of Linguistic Skill in Twins, Singletons and Siblings, and only Children from age five to ten years : By EDITH A. DAVIS. (pp. 165. 9s.)

The Effects of Praise and Competition on the Persisting Behavior of Kindergarten Children : By THETA HOLMES WOLF. (pp. 138. 9s.)
All published by the University of Minnesota Press, Minneapolis, Humphrey Milford, and Oxford University Press,

These three books constitute the *Child Welfare Monographs*, Nos. 13, 14 and 15 respectively, edited by the Director of Minnesota Institute of Child Welfare, Dr. John E. Anderson, who writes a brief foreword to each. Like others of this series they are characterised by an evident desire, first, to establish clear facts about child development, and by an absence of any selecting of facts to fit a particular theory, and a ready recognition of the limitations of the particular investigation.

The enquiry as to children's concepts of causal relations deals with 732 children between the ages of eight and sixteen. Questions were put to the children on such topics as follows : What makes the wind blow ? What makes the snow ? Why do boats float on water ? What causes thunder ? etc., and a number more decidedly of a type that might occur in early school work in general science. It is important to bear in mind the type of questions put because it no doubt explains several of the main findings, more particularly the surprising result that there was only a small connection between the scores and estimated intelligence, within any given age group ; whereas the relation between score and school grade (age being held constant) was considerably higher. As the author points out : " This suggests that the answers to these questions are more directly determined by factors involved in school experience than they are by intelligence."

The author emphasises the great influence which the particular content of the question has on the type of the answers given and concludes that results contradict some of the main findings of Piaget on this topic. A final point of interest is that the small group of kindergarten children show surprisingly high scores.

The second study, dealing with linguistic skill, was based upon experiments with over 400 children of whom 166 were twins and 97 only children. Due attention was paid to the distribution of the children in different social grades, and also to the length of schooling. Various interesting toys and " play situations " were arranged and the children were encouraged to talk about these. Their remarks form the substance of the material for analysis, for an estimate of talkativeness, and so forth.

The main conclusions were as follows :

- (1) Only children are definitely superior to children with siblings in every phase of linguistic skill (even allowing for the greater frequency of only children in upper social classes).
- (2) Singletons with siblings are in turn somewhat superior to twins.
- (3) Twins from the upper occupational groups by 9½ years have practically overcome their language handicap, but twins from the lower occupational groups have made relatively little progress.
- (4) Twins of the ages studied in this investigation are especially retarded in perfection of articulation. The inferiority is marked during the kindergarten period, particularly in twins from the lower occupational groups.

The third of these books deals only with five-year-old children in the kindergarten school. The tasks given them were the setting up of various toy models, a game of toss with quoits, putting pins into a peg board with a pair of tweezers, an automatic, repetitive task, namely, a cancellation game in which, in pages of

figures, all the dolls and only the dolls had to be crossed out. Finally there was the dull task of dropping little lead balls into a box through a small hole. Three different types of incentives were used ; the first was simply that intrinsic in the material itself ; secondly, praise was given for success ; third, there was a competitive situation set up, four children being picked at a time and urged to compete with one another. With regard to this last the author wisely points out that it is "impossible to know how much of the increased performance should be attributed to competition and how much to mere sociability." For four of the five tasks the effectiveness of the stimulus increased in the following order : first the task without added incentive ; second, praise by the experimenter ; third, competition. Praise was considerably more effective with the quoits than with the other three tasks. Competition showed a greater influence relatively on this very game of quoits. It also specially affected the two simple repetitive tasks of cancellation and dropping balls into the box. The author concludes then that the effectiveness of incentives is limited by the nature of the tasks performed. One outstanding example was the difficult construction problem of the "tinker toy." "The results of this study suggest that persisting behavior is both general and specific. There is a low but positive relationship between persisting performances in different tasks and different incentive situations."

C.W.V.

Introduction to Experimental Psychology (third edition, enlarged and revised) : By C. W. VALENTINE. (London : University Tutorial Press, Ltd., pp. x + 283, 4s. 6d.)

Since Prof. Valentine's manual was first published, the introduction of a course of practical psychology into the curriculum for students who are being trained as teachers has become far more general. In reviewing the previous edition in *The Forum of Education* it was pointed out how seriously such proposals were handicapped by the lack of an appropriate textbook of experimental exercises. What is needed is a cheap and handy volume well within the limited means of the elementary student : the exercises should be suitable for large classes carrying out the experiments in groups of twenty or thirty at a time ; they should involve no costly apparatus, and pre-suppose no laboratory expensively equipped. This book is the only publication in this country which fulfils these requirements ; what is more, it fulfils them so well that no other book has been needed or attempted. That after nearly a quarter of a century the demand for it is still so great as to justify a third edition is, of itself, a high testimony to the merits of the work.

In the meantime, psychology has progressed. And in no other field has it advanced so rapidly as in its application to educational problems. The new edition comes into step with these changes. Fresh chapters have been added, dealing more especially with the topics of perception, apperception and imagination, accuracy of report and susceptibility to suggestion, and the intuitive judgment of character and intelligence. Many of the older experiments and tests have been revised or supplemented in the light of recent work on backward and subnormal children, and admirably bring home to the beginner the value of more scientific methods in studying such cases. To help the teacher to evaluate his own tests and examinations, experiments have been inserted to explain the nature of correlation and its use in schools ; and the discussion on general intelligence and group factors has been brought carefully up to date. The chapters on attention and memory have been revised and extended : the latter now includes new experiments suggested by Prof. Bartlett's book on *Remembering*. The invasion of psychological methods into unexpected fields is illustrated by experiments on the appreciation of colour, poetry, pictures and literary style.

The book as a whole is at once so sound and comprehensive that it should be of value, not merely to the intending teacher, but also to the student of general psychology who needs an elementary introduction, and particularly to those who are studying it at home or as external students without access to laboratory facilities or a trained instructor. In a subject where so much is published which the beginner need not and should not read, the selected references in the footnotes and the classified bibliography at the end will be found particularly helpful. Even the

advanced research worker may profit by the numerous original and suggestive comments scattered through the pages (e.g., the much needed advice "to approach the study of mental types rather by analysing the more intellectual material of ideas than by the study of the vaguer temperaments") and by the concise summaries of results reached by Prof. Valentine himself in important educational fields.

The teacher of psychology will benefit most of all. With a sure flair for what is soundest and most suggestive in modern psychological theory, Prof. Valentine combines a long experience of the practical needs of the elementary student. Every instructor, therefore, who has to devise an experimental course in the subject will be particularly grateful to him for thus placing both his knowledge and his experience at their disposal. C.B.

Human Nature at Work: By JEAN L. SHEPARD. (Sir Isaac Pitman and Sons, Ltd., 1938, pp. xi+219. 7s. 6d. net.)

The American people has now to a large extent completed the conquest of the physical features of its territory, and its pioneering energy is pushing ever more strongly forward along a front where even greater triumphs are to be won. The problem of how people are to live and work together happily in groups is receiving more and more attention, not merely by psychologists but also by business men, who, in America at least, have proved to their own satisfaction that its successful solution brings a return in dollars and cents.

This book is an excellent example of the propaganda now being devoted to the improvement of human relations in business. The author claims that she, along with other personnel workers in America, has gone far beyond European ideas, which tend to represent the individual worker as a "physical entity that needs care like a machine," to the study of emotional factors in human beings at work. It is these factors more than any others which determine the success or failure of an individual, and the author develops this theme in a very readable series of case histories supported by general observations on the principles involved.

One short extract will serve to indicate how the author can give a stimulating interpretation of the simplest of everyday situations:

"One executive said he always sat at his desk and wondered 'why the people on the other side were so ill at ease, although he did nothing consciously to put them at a disadvantage. In truth, his manner was irreproachable. . . . One day he happened to be engaged in the centre of his office and found that, without his position behind the desk, that symbol of authority, he was not himself. Something was missing from his usual comfort and security."

"This interested the executive so much that he decided to try the experiment of giving the other fellow a chance on an equal footing. He remodelled his office, making it a simple living-room with a small sofa, comfortable chairs, tables placed for convenience, but no desk. He met everyone directly as man to man, and was gratified at the spontaneous expression and change of ideas that resulted."

Similar insight is brought to bear on a large variety of other problems, from the interviewing of applicants to the treatment of older employees whose ways are set; and although the book may be read once for sheer interest, it should be read several times by anyone set in authority, to ensure that the maximum profit is obtained from the sympathetic discussion of so many human difficulties.

E.P.H.

The Riddle of Life: A Survey of Theories: By WILLIAM McDougall. (Methuen, 1938, pp. 279. 7s. 6d. net.)

It is sad to think that this is the last book McDougall lived to write, but it is a fitting coping stone to his teaching during almost fifty years. For in it he gives his final conclusions about the difference between living and material things, a question which had been the theme of his earliest works and with which all his writing had been more or less connected.

Most of the book is devoted to a critical summary of the chief answers which have been given to this question in recent years. He discusses briefly but incisively mechanical biology, emergent evolution, and holism, insisting everywhere on the reality and importance of psychical teleological activity. "The psychical factor in the life of organisms is of their very essence, and throughout the scale of organic evolution it has become of increasing efficiency" (pp. 236). The conception of a living organism which he himself prefers is that of a hierarchy of monads in which physical and psychical activities, though connected with each other, are radically different. His account of monadic and dualistic theories is therefore more sympathetic.

In spite of the physical disabilities under which the book must have been written we find the clarity and vigour of statement, the wide knowledge and flashes of insight, and whole-hearted pursuit of truth which we have learned to expect from his other works. If the limits of his space and a certain lack of philosophical rigour lead us to feel at times that there is more to be said than he implies, he has, at any rate, given us an illuminating introduction to the study of a problem which is fundamentally important in biology, psychology, and philosophy. His masterly statement of his own views and broad survey of other theories will serve as a valuable guide to readers making their first acquaintance with a field which is being actively explored, and will be welcomed by those to whom his main conclusions are familiar.

H.B.S.

Studies in Clinical Psychology: Edited by LEE EDWARD TRAVIS; and *University of Iowa Studies in Psychology No. 22*: Edited by CHRISTIAN A. RUCKMICK. (Psychological Review Co., Princetown, N.J., pp. 250. \$6 30 cents.)

This volume includes a number of doctoral theses involving an analysis of stuttering and the neurophysiological processes underlying voluntary conditioned reflex and reflex behaviour in man. The data seem to warrant the following conclusions, among others: (1) a decrease in efficiency of eye movements in stutterers during stuttering seems to be a reaction to difficulty present in one of three physiological mechanisms used in the speaking act, viz., verbalization; (2) two general forces—the reaction to social confrontation, and the ontogenetic factors surrounding the malady of stuttering interact with the neurophysiologically susceptible organism of the stutterer and produce stuttering in those situations where the three are present to a sufficient degree; (3) susceptibility of the stutterer to changes in the speaking situation is suggestive of the efficacy of a therapy which will increase his ability in social confrontation and develop within him a more objective point of view towards his disorder.

Psychology and the Religious Quest: By Dr. R. B. CATTELL. (Nelson, pp. 190. 2s. Od.)

One of the aims of this book is to show that certain scientists who have, to their own satisfaction, exploded religion do not go far enough. Dr. Cattell believes that the logical outcome of the theory of the group mind is something to which men in the past have given the name God, but which he calls the Theopsyché. He traces the relation of man to this in the field of morals and ethics, and insists that this consideration is urgent to-day.

His dismissal of present-day religious beliefs is confidently dogmatic, but his remarks on some of the values of religion are sympathetic. He gives a brief but interesting survey of the evolution of certain religious beliefs as shown by anthropologists.

Many who do not agree with his position will do well to consider his plea that the psychological investigation of religious problems, especially among youth, is a pressing need to-day. His sympathetic insistence that religion needs to be considered scientifically gives much value to this book.

T.A.G.

The London Head Teachers' Association, 1888-1938. (University of London Press, pp. 143. 3s. 6d. net.)

The London County Council celebrates its own jubilee with proper pomp this year. The London Head Teachers' Association came into being a year before the L.C.C., and has celebrated its fifty years of existence by the publication of a jubilee souvenir book. There are something like fifty illustrations, old and new, ranging from the infant galleries, the massed classes and massed drills, the flowing robes and fierce moustaches of the old-time staffs, to the open-air agility, the health services, practical activities, physical education and unobtrusive teachers of to-day. The period covered by the record, half of it pre-war, saw "the transformation of the whole London working population from a semi-educated to an educated community." When the L.H.T.A. began, payment by results was still in operation, a superannuation scheme was still to come, and, as Mr. Rich's foreword notes, teachers had only recently ceased to be stoned in the streets as they went to and from their schools.

The Secretary, the President and others survey the half century of the Association's activities, and the book is fortunate in its guest-artist contributions. The Education Officer writes a sympathetic foreword, the Leader of the Council (Mr. Herbert Morrison) pays tribute to the elementary school, and four racy chapters are contributed by well-known officers, past or present, in the Council's educational service. Mr. George Sampson has some wise and mellowed reminiscences of pupil-teacher times, with a side glance at young men and women of to-day who "sometimes go gracefully through the motions of teaching before their classes and feel that their responsibility is thus discharged." Dr. Ballard notes from personal experience that during the fifty years the normal size of classes has been halved. Professor Burt maintains that education is quickly turning from an art into a science and that the teacher, who in the first half of the period under review had become more human, has in the second half of it become more scientific, without any sacrifice of his humanity. "The old teacher ruled his pupils; the new teacher studies them." Dr. F. H. Spencer writes in lively vein about inspectors, both the H.M.I. and the L.E.A. type, having himself had experience of both kinds. He is "not at all convinced that to leave teaching for inspection is calculated to increase the happiness or the comfort of those who make the change." An inspector has to be "a good arbitrator, a good diplomat, a good speaker, and, in general, a kindly, resolute, intelligent, imperturbable archangel."

The special chapters bring out the fact that there are now nearly as many pupils in the L.C.C. Selective Central Schools as in the public Secondary schools of London. Technical education, the evening institutes and the day continuation schools are briefly dealt with, though there is no reference to the Literary Institutes—a special feature of the London system. In general, however, the book is a comprehensive and very readable record of "fifty glorious years" of educational service.

A.E.D.

The Nineteen Thirty-Eight Mental Measurements Yearbook. Edited by O. K. BUROS. (New Brunswick: Rutgers University Press, pp. 415, \$3.)

For several years Professor Buros has been regarded as one of the best friends of psychologists engaged in measurement. His *Educational, Psychological and Personality Tests of 1933, 1934 and 1935* is a bibliography of eighty-three pages. His corresponding volume for 1936, containing critical reviews of tests by experts, runs to 141 pages. Now, however, the tremendous activity of test technicians and writers of books on mental measurements has made it necessary to expand the scope of the service to the extent of 415 pages. According to a conservative estimate there are over two hundred psychological tests standardised every year in the United States alone, and as this Yearbook also reviews books on mental measurements, research and statistical methodology, etc., for the whole of the English-speaking world, it is clear that the task of keeping abreast of the times without its aid is almost hopeless.

On the front of the jacket appear the names of 135 expert test reviewers from the universities and colleges of the United States whose co-operation was enlisted. Review excerpts from 185 journals in English-speaking countries are presented so as to assist readers in making selections from among the hundreds of books and tests published. One outcome of this Yearbook will be to impel test authors and publishers "to construct fewer and better tests and to furnish a great deal more information concerning the construction, validation, use, and limitations of their tests." As the editor points out: "Test users have every right to demand that test authors and publishers present full particulars concerning the methods used in constructing and validating the tests which they place on the market."

For each test the following particulars are given: title, author, publisher, group for which it was constructed, date, cost, working time, and references. Already the editor has in view still further improvements for the 1939 Yearbook, such as a more adequate representation among the reviewers of able classroom teachers.

The venture deserves every encouragement and it is fervently to be hoped that the requisite finances will long be available to enable Professor Buross and his staff to continue these ambitious and helpful activities. J.L.W.J.

The Industrial Worker: By T. N. WHITEHEAD. (Oxford University Press. Vol. I and II. 21s.)

This is in many respects a monumental book. It represents twelve years' work, during six of which a group of five girls worked under the closest observation in an experimental test room devised and maintained by the Western Electric Company at Hawthorne. Except that the girls were carrying out a very routine industrial assembly job, the conditions of work were unlike those in modern industry, and approximated to those of the small family industry of earlier times.

The part of the experiment that appealed to the girls most was the fact that "We was told to work as we feel and we do," but as every teacher knows, there are limits to this attitude. In this small group of five two Polish girls badly upset the whole social atmosphere. Output suffered. They were removed at the end of the year.

In their place came two friends, one Italian, a dominant personality. She realised that the experiment, which was under the eye of the Executives, was not coming up to expectations, and soon discovered subtle ways of speeding up the group. Output increased and everyone was pleased, but the pace got so hot that the leader herself exclaimed: "I always get so nervous that I cannot stand it. Everything must be done in a hurry. . . ."

The inevitable reaction that followed was accentuated by the fact that management's interest had turned towards a new enquiry. The leader was disheartened, the girls became irritable, complained of the ventilation, their supervisor, and were even suspicious of the accuracy of their pay.

The last part of the book shows by an elaborate statistical analysis of the records and comparisons with the remarks of the girls, that friendships and hates operate in industry in much the same way that they do in school. Little social groups form, break up, and individual work is influenced accordingly.

With such a small group doing work in which boredom was the greatest enemy, social factors were bound to dominate the situation and North Whitehead has done heroic work in rescuing the facts from a prodigious welter of records.

G.H.M.

Psychology of Music: By CARL E. SEASHORE. (McGraw-Hill, pp. 408. 24s. 0d.)

This work is a retrospective synthesis of the results of forty years' research in the psychology of music, a research conducted in laboratories equipped with every modern scientific device for the accurate measurement, recording and graphical representation of the most minute modifications in pitch, intensity and timbre throughout the duration of any sound, a research conducted by a body of enthusiastic experts, inspired by the direction of Professor Seashore. It is obvious, therefore, that such a work must be regarded as a landmark in the history of the psychology

of music. A fundamental hypothesis is that "the medium of musical art lies primarily in artistic deviation from the fixed and regular," and that, therefore, musical performance is capable of quantitative measurement in terms of deviation from the "true" in each and all of the four attributes of sound: pitch, intensity, duration and timbre. In addition to extensive records of such deviations, the work includes an interesting analysis of the quality of timbre, and differentiates the timbre of all the main orchestral instruments by contrasting their timbre spectra; the timbre of the bassoon is treated in detail. Further, the whole range of musical ability, innate and acquired, is discussed, with factors implied in training, and a chapter is added on musical aesthetics. The work is generously illustrated, includes a large number of statistical records, and an extensive bibliography.

To those who have opposed the extravagant claims of the empiricists the work contains a welcome expression of the limitations of the scientific and objective approach to the study of what constitutes goodness in music. A science of music is essentially a science of musical performance, superimposed on the creative work of the musician, and the merits of a particular work are clearly distinct from those of a particular interpretation of it. Only the latter is capable of exact measurement; the former, like the beauty of a poem, is an elusive value outside the domain of the scientist.

J.M.

Play in the Infants' School: By E. R. BOYCE. (Methuen, pp. 185. 5s.)

The reorganisation of a school is an enterprise demanding skill and courage. Miss Boyd's account of her experiment in the Raleigh Infants' School shows that she lacked neither.

The children in this East End school suffered from the physical ills and the mental stunting common to small slum dwellers. The existing school curriculum failed to meet the real needs of the children and Miss Boyce, with good reason, rejected the "project" method. So the children were given freedom to develop their abilities and to acquire mental and physical poise through their self-directed activities in an environment designed to compensate, in some degree, for the deficiencies of their homes. The experiment is important as a considered attempt to meet the special needs of a particular group of children.

Miss Boyce's description of their activities—shop-keeping, cooking, construction, imaginative play, and their delightful and vigorous performances in the school theatre—conveys the impression of their eagerness and concentration, and shows how wide an experience can be gained through play. Moreover, it provides additional material for the study of the play of young children.

In the last section of the book the introduction of formal learning is described and the attainments of the children are frankly discussed. The reports suggest that in speech, dramatic ability, and mathematical skill they surpassed children from similar homes in schools of a more orthodox type. Their eagerness to learn and their wider background would compensate for deficiencies of achievement in reading.

An Educational Failure: a School Inspector's Story: By F. H. HAYWARD. (Duckworth, 1938, pp. xvi+379, 7s. 6d. net.)

Dr. Hayward has long been known as an enthusiastic advocate of school celebrations, which, he believes, might revolutionise our ethical and religious teaching. It must be admitted that his drastic criticism of the kind of teaching now given in some schools and churches is not without justification, though his strictures are far too sweeping. We may also agree the type of celebration he recommends has been unduly neglected. Whether it could achieve all the results he claims is much more doubtful, but if he had given us a reasoned account of its theoretical basis and systematic suggestions for experimenting with it, his book might have been suggestive and practically useful. Unfortunately it is concerned mainly with his failure to induce the London County Council to organise school celebrations on a comprehensive scale. It is written under the influence of his sense of disappointment, includes much that is only of personal or passing interest, and cannot be said to make any serious contribution to the discussion of a movement with attractive potentialities.

H.B.S.

Your Life's Work: By E. D. LABORDE. (London: Thornton Butterworth, 1939, pp. 223. 5s.)

Some years ago Dr. Laborde, who is a careers master, wrote a book on occupations, which was probably found by many people to be a useful one. Unfortunately, the same cannot be said of this production. From its first paragraph—which seems to defy grammatical analysis—to its last, it is blemished by dogmatism, misconception, mis-statement and a number of other qualities which are perhaps even less desirable in the writings of careers advisers than in those of ordinary mortals. Several of these characteristics can be discerned without difficulty in the brief passage (on page 154) in which Dr. Laborde deals with the psychologists' contribution to the problem with which he himself is primarily concerned.

"In recent years there have come into existence a number of methods of vocational testing, which claim to treat the boy as a subject and at the end of a series of processes to announce the careers for which his bent and talents fit him. Such tests are widely used in firms whose recruits are unable to produce a house-master's summary of character. As, however, they are largely formed on the collated estimates of character as given by parents and schoolmasters, they are unnecessary when the schoolmasters are capable of giving their own judgments."

A.R.

Alcohol and Human Life: By Dr. COURTENAY C. WEEKS, with Foreword by Sir Thomas Barlow, Bart. (H. K. Lewis and Co., Ltd., 1938, pp. 454. 6s.)

This work is partly a revision of *Alcohol and the Human Body*, by the late Sir Victor Horsley, the late Dr. Mary Sturge, and others. In the opinion of the author alcoholic indulgence is still the greatest single enemy that Britain has to fear, the one which "allies itself with all that mars and spoils human life," and an impressive collection of social and scientific data is assembled here to support that belief. An important chapter is the one dealing with alcohol and motor driving, in which the effects of small doses of alcohol on motoring efficiency are discussed. These effects show themselves in diminished powers of attention, concentration and sense of anticipation and responsibility. There is interference with central and peripheral vision, and there is also an inaccurate and slowed response to emergency impressions.

In view of the fact that owing to the sense of well-being caused by alcohol there is an uncritical satisfaction with these faulty psycho-physiological conditions, total abstinence before and during motoring and flying appears to be essential.

Sex, Friendship and Marriage: By K. C. BARNES and G. F. BARNES. (George Allen and Unwin, pp. 199. 6s. net.)

This book gives a popular discussion of various problems of love, courtship and marriage, together with a more technical appendix dealing with the physiology and anatomy of sex development. The discussion is on the whole marked by good sense, and would be useful reading for a good many young men and women, although I doubt its value for older boys and girls, which is suggested on the back cover. Furthermore the technical nature of the appendix, which occupies about one-third of the book, lessens the unity of the book, and some parts of the appendix seem to be quite unnecessary for those who would clearly be interested in the main part of the book.

High, Wide and Deep: By C. MADELEINE DIXON. (George Allen and Unwin, pp. 300. 12s. 6d. net.)

The sub-title of this book, *Discovering the Pre-School Child*, indicates broadly its purpose. It is a somewhat diffuse description of work and play in the nursery school, with a good deal of discussion, and some popular and not too well established psychological interpretations. It will, no doubt, be quite suggestive to teachers of nursery school and infant pupils.

THE BEARING OF GENERAL AND SPECIAL
ABILITIES UPON SCHOLASTIC SUCCESS AT
THE BEGINNING AND END OF A SECONDARY
SCHOOL CAREER.

BY MARY ORMISTON.

PART II : SENIOR CHILDREN.

- III.—(ii) *The Senior Group.*
The outside criteria.
The tests used.
The analysis of the School Certificate results.
The discussion of these results for the subjects
analysed.
IV.—*Summary and conclusions.*

(ii) *The Senior Group.*

The outside criteria.—The results of the School Certificate examination were taken as the criteria of success at the end of the secondary school career. Although this had many advantages in that a common test was applied to many schools, and that the marking throughout was on the same carefully standardised basis, and that there was no chance of the personal element entering to favour any child, it yet presented difficulties. The most obvious was the probable presence of a teaching factor. In this examination not only would the standard of tuition vary from school to school, but also in the same school from one subject to another. To obviate any possible factor arising from such differences in teaching, the correlations were obtained separately for each subject for each form.¹ A further difficulty was met in that for each school the same children did not take each subject in the examination, thus some would take Chemistry, others Botany. Where the numbers were small the subject was omitted in the final analysis.

The tests used.—In the preliminary work, tests had been made for the children at the end of their secondary school career which were as comparable as possible to the tests used for the junior children. Similar fundamentals were involved in both batteries of tests and in the analysis

¹ FISHER : *Statistical Methods for Research Workers*, p. 170.

of the two sets of preliminary results the factor make-up of each respective pair of tests was similar. Thus, for example, for both the younger children and those of School Certificate standing the respective Verbal Analogies Test was most highly loaded with g .

Unfortunately the number of tests had to be reduced to fourteen, as a shorter time was available for testing. Thus the individual Passalong Test had to be dropped, and the Logical Selection, Dot Pattern Perception, and Overlapping Shapes Tests were also omitted. The German Test was given, but all the results were not included in the analysis, as it was found that some of the children who had a knowledge of the language had an unfair advantage.

An analysis of the results of the tests alone was first made to see that they were working in the way anticipated. As in the junior tests, so here the test of Verbal Analogies proved to be most significantly loaded with the g factor (the loading was $\cdot 713$). Considering the five years difference in age between the groups, it was highly satisfying to have found a test so similarly loaded with g . The Finnish Test, which had for the juniors been taken as that most highly loaded with the v factor, proved to be similarly highly loaded in this result, the loading being $\cdot 522$, next in order being the Opposites Test with a loading of $\cdot 316$. The Mechanical Arithmetic Test again proved to be that most highly loaded with factor n , the loading being $\cdot 575$. The Paper Formboard Test was, as in the junior group, highly loaded with the spatial factor ($\cdot 432$), and the Spatial Imagery Test with the factor termed S_2 , and tentatively interpreted as a spatial imagery factor.

These tests, namely Verbal Analogies, Finnish, Mechanical Arithmetic, Paper Formboard, and Paper Folding, had been used for the axis of the g , v , n , S_1 and S_2 factors respectively in this analysis of the senior work as in the junior group. The loadings in the parallel tests given to these older children were similarly high in each of the respective factors as were those of the younger group. The analysis of success at the beginning of the secondary school career had been made in terms of the factors g , v , n , S_1 and S_2 . If then an analysis of success at the end of the secondary school career were made in terms of these same factors, the rotations for which were made through parallel tests to those used for the rotations in the junior group, the results of the two analyses would be comparable.

The analysis of the School Certificate results.—The actual failing or passing the School Certificate examination was ignored, the actual number of marks obtained in each paper being taken as the criterion of success. As the factors emerging after continuing the analysis necessarily

become smaller and, therefore, the probable errors proportionately larger, the first three factors were found to be of greatest significance. If all the School Certificate subjects had been analysed at one and the same time, owing to the multiplicity of factors involved, only the largest factors obtained would have been significant. Therefore, the School Certificate subjects were subdivided into groups.

The Language Group.—The first group analysed included English Language, English Literature, History, French, and Latin. The tests used for the analysis of this group were Non-Verbal Analogies, Verbal Analogies, Finnish, and German.

The Non-Verbal Analogies Test was used for the g rotation in order to escape any verbal tendency. This g factor, as was of course to be expected, played an important part in each of the subjects, the loadings being :

	g
English Language	·685
English Literature	·544
History	·424
French	·762
Latin	·670

The verbal factor as obtained through the Finnish Test was of importance in English Language, the loading being ·284, but not in the other subjects. The swing away from the compositional type of question in History, French, Latin, and even in English Literature, which has taken place in the type of question set in the School Certificate papers, seems the probable cause of this lack of the v factor. Its presence in the English Composition would agree with the results of the junior analyses and with the above suggestion. A further factor, termed y , not obtained in the tests, was common to Latin (·135) and French (·140); and another factor, termed x , only found in Verbal Analogies, was operating in History (·438), English Literature (·480), and English Language (·299).

The Number Group.—Fortunately, the School Certificate marks were available separately for Algebra, Arithmetic, and Geometry. These marks were now analysed, together with those obtained in the tests for Mechanical and Non-Mechanical Arithmetic and Verbal Analogies. The first rotation was made through Verbal Analogies, the second rotation

was made through Mechanical Arithmetic and the resulting g and n loadings obtained are given below. The third factor, before rotation, was greatest in Algebra, and not present in any of the Psychological Tests, and accordingly rotation was made through this subject, the resulting factor being termed α .

	g	n	α
Arithmetic	·573	·377	·229
Algebra.....	·214	·643	·371
Geometry	·275	·333	·695

A third grouping of some of the School Certificate subjects was made and termed the—

Spatial Group.—The examination marks in Geography, Geometry, Botany, and Art were analysed by means of scores obtained in the Verbal Analogies, Paper Formboard, Paper Folding and Spatial Imagery tests. The first rotation for the g factor was made through the Verbal Analogies Test. The second and third rotations were made through the Paper Folding and Paper Formboard tests, and the following loadings resulted :

	g	S_1	S_2
Geography	·311	·419	·420
Geometry.....	·275	·330	·711
Botany	·230	·507	·006
Art.....	—·266	·756	·150

A discussion of the factor loadings obtained for the School Certificate subjects analysed.—From the preceding results some suggestions may be made with regard to the contribution of general and special abilities to success in the School Certificate subjects analysed in this research.

The English Examination.

(a) *The Language Paper.*—The first compulsory question was concerned with Comprehension, and was answered by a series of short

answers. The second was English Composition. The other questions involved grammatical knowledge and comprehension. Restating the factorial loadings of the analysis, namely, $g \cdot 685$, $v \cdot 284$, $\alpha \cdot 299$, it can be seen that in terms of these factors success depended primarily upon the general factor, to some smaller extent upon the verbal factor; and the factor termed α , which, it is suggested, is a factor involving character or attitude to work¹. This result is in keeping with that obtained in the analysis of the English Comprehension and Composition papers for the junior group in District B. In the latter it was found that in the Comprehension section of the examination the g factor was dominant, while in the Composition the v factor was to the fore. In the School Certificate paper no separate marks were available for the different sections, but since the Comprehension aspect of the paper is greater than the Compositional, the importance of the g factor was to be expected.

(b) *The Literature Paper.*—The factor loadings found for this paper were: $g \cdot 544$, $v \cdot 046$, and $\alpha \cdot 480$. The English Literature Paper naturally involved many questions calling for knowledge of the work studied. Some questions involving criticism of extracts quoted were also included. The g factor is again of first importance, but the α factor is ranking second. As suggested above, this factor seems to be one involving character or attitude to work. The Literature Paper would give more scope for such a factor than the Language Paper. It can be concluded that success in the English Examination as a whole depended upon g , α , and v , in that order of importance.

The History Examination Paper.—The loadings obtained were: $g \cdot 424$, $v \cdot 182$, and $\alpha \cdot 438$. The possession of knowledge seemed the most essential factor in answering this paper, and the scope for composition was limited. It is not then surprising to find that the g factor and that termed α , involving attitude to work, should be of most importance.

The French Examination Paper.—The loadings obtained were: $g \cdot 762$, $v \cdot 042$, $y \cdot 140$. The French Paper consisted chiefly of translations from English to French and *vice versa*, some Comprehension

¹ This suggestion is made on the grounds that α is found in the examination papers yet not in the tests, that in the preliminary work (see page 165) a trait estimated by the teachers as 'persistence' had been noted and that such a factor had been suggested by the work of:

ALEXANDER: *Brit. Jour. Psych.*, Mon. Supp. xix.

WILSON: *Brit. Jour. Educ. Psych.*, vol. iii, 1933.

FLANAGAN: *Factor Analysis in the Study of Personality*.

For further discussion see page 220.

questions and a short Dictation. The importance of the *g* factor for success in this subject is high and the factor termed 'verbal,' which has been repeatedly found to operate in the English Composition, is not present in this subject. A new factor termed *y*, which was not found in any of the tests, but which occurred also in Latin, is the other factor to be considered. This factor *y* is not present in the English Language nor Literature Paper, nor yet in the tests involving memory of words. It may then be some special factor which is involved in translations to and from a foreign language.¹

The Latin Examination Paper.—The loadings obtained were: *g* .670, *v* —.349, and *y* .135. The similarity between the factors involved in this and the French Paper are indeed striking, and what has been said for the French Paper applies similarly to this paper.

The absence of the so termed 'verbal' factor is marked and it is interesting to note the correlations found by Earle¹ between English and Latin and Verbal Tests of Opposites and Synonyms, and Mixed Sentences.

<i>Correlations after Earle.</i>	<i>English.</i>	<i>Latin.</i>
Opposite and Synonyms.....	.37	— .05
Mixed Sentences21	— .32

In the analysis of this Latin Examination Paper the Opposites Test has been found to be loaded with the verbal factor to the extent of .316, while in Latin the *v* loading was —.349.

The Arithmetic Examination Paper.—The loadings were: *g* .573, *n* .377, and *x* .299. In Section A of the Arithmetic Paper some mechanical questions were involved, but Section B was entirely composed of problems. The analysis of the Arithmetic of the junior group, where the marks for the Mechanical and Problem sections were separately available, would lead to the conclusion that the *g* factor would be dominant in the Problem section of this School Certificate examination, and hence the importance of this factor, while the *n* factor would be of greatest importance in the Mechanical Section A. Again, the presence of the *x*, or the factor involving attitude to work, is considerable.

¹ Although the factor *x* is not one of the three first extracted in this analysis, it should not be inferred that *x* is not operating in this subject for, on further analysis, it might be found.

The Algebra Examination Paper.—The loadings obtained were : g .214, n .643, and x .371. Only four of the fourteen questions set in this paper were problems. As is seen above, the n factor contributed most towards success in this subject, followed by the x and g factors. As much mechanical number work was involved, and little of the problem type, the importance of the n factor is understandable.

The Geometry Examination Paper.—The loadings obtained were : g .275, n .333, S_1 .330, S_2 .711, and x .695.¹

The Geometry Paper was in two sections, A and B. In Section A, the first part of three out of every four questions was composed of instructions to construct figures to given directions. The second part of each question involved some simple deduction. The six questions in Section B each involved the exposition of some theorem and a rider thereon. The learning of theorems, the understanding of the same, the formulating and carrying out of constructions are all playing their part in this examination. Such complexity in the processes here involved would reasonably be expected to depend upon many factors.

The importance of the attitude to work, and spatial and spatial imagery factors are dominant in this examination, followed by the number and the general factor.

The Geography Examination Paper.—The loadings obtained were : g .311, S_1 .419, and S_2 .420. In each question map work was involved and a considerable knowledge of facts. That the factor involving spatial imagery, and the factor involving ability to deal with spatial relations, should be important in this subject was to be expected, for they would be playing their part not only in the map work but also in the realisation of relationships in space involved in the factual answers. The g factor was, of course, playing its part as in other subjects.

The Botany Examination Paper.—The loadings obtained were : g .230, S_1 .507, while S_2 was almost negligible, being .006.

In this paper the use of sketches is frequently demanded to illustrate the written answers, or, as in other cases, the examinee is required to describe and explain a given sketch. An ability to deal with spatial relations could then be expected to be playing an important part. The factor of spatial imagery is insignificant. The g factor is of secondary importance.

¹ The Geometry results were analysed in both the number and spatial groups of subjects. As a result of the indeterminateness of the tests used for rotation and the approximations made in calculations, some inexactness would inevitably creep into the resultant loadings, thus $\sum h^2$ is slightly greater than 1.

The Art Examination.—The loadings obtained were: g $\cdot 266$, S_1 $\cdot 756$, and S_2 $\cdot 150$. In this investigation, Art is shown to be independent of g , but highly saturated with factor S_1 , and to some extent dependent upon S_2 . The factors S_1 and S_2 are not dependent upon the *ability to draw*, for the tests through which these factors were detected did not depend upon ability to draw. This result is similar to that found by Flanagan¹ and by Alexander.²

It is unfortunate that owing to the amount of testing time which would have been required from the School Certificate group, tests for Logical Reasoning and Persistence were not included. Further, although some data were obtained for Physics or Chemistry, the numbers involved in each form were as a rule small, and no conclusions could be drawn with safety.

It can be seen that the bearing of both general and special abilities upon success at the end of the secondary school career varies enormously from one subject to another.

Before summing up, some limitations of the method employed must be noted. Throughout, and especially in the preliminary work, an endeavour was made to obtain as pure a test as possible whenever a rotation was necessary. Thus Verbal Analogies, a test which was found to be highly loaded with g factor ($\cdot 886$) leaving little scope for even a v factor, was used for the rotation of the g axis, but even so small a v loading as was possibly involved would lead to a slight indeterminateness in the location of the axis. It can be claimed that such indeterminateness was small for this g factor. When the factors S_1 and S_2 were obtained, the position was not so clear cut, for the tests involving the so-called 'spatial' factor tended also to involve that termed 'spatial imagery.' The Paper Formboard Test was used for S_1 , as it had been found to have but a small specific link with the tests highly saturated with spatial imagery, nevertheless some indeterminateness must be admitted. Since no previously analysed test saturated with the α factor was available, the figures obtained for the loadings in this factor had to be obtained by rotation through some school subject. That this α factor could be one involving either speed, age, difference in teaching, or memory, had been disproved by the preliminary work, and by the method of obtaining the correlations for each form separately. Since it is none of these, and

¹ FLANAGAN : *Factor Analysis in the Study of Personality*.

² ALEXANDER : *Brit. Jour. Psych.*, Mon. Supp. xix.

is common to school subjects yet not to tests, it seems highly probable that attitude to work could be an acceptable interpretation. The work of other investigators¹ would also support this suggestion.

Considering the general success at the end of the secondary school career as being the sum of the successes analysed, the following summary can be made.

The g factor is by far the most important contributor to success, the summed loadings being almost twice as great as those for any other factor, it being found in all the subjects considered, excepting Art. The x factor, namely that involving character, is second in importance, and it seems probable that it is playing a part in most of the subjects considered. The factor involving the ability to deal with spatial relationships ranks third, followed by the number, spatial imagery, and verbal factors, and by the factor termed y , which was involved in French and Latin.

Although, owing to the indeterminateness involved, little absolute value can be attached to the numerical summation of the loadings, their relative value may be of some interest. Such approximate values as obtained from the analyses considered are given in the following table.

TABLE SHOWING THE RELATIVE IMPORTANCE OF THE FACTORS MAKING FOR SUCCESS IN THE SCHOOL CERTIFICATE EXAMINATION AS ANALYSED IN THIS RESEARCH.

g	The general factor.....	17
x	A character factor involving persistence and attitude to work	9
S_1	A factor involving the ability to deal with spatial relations	7
n	A number factor	5
S_2	A factor involving spatial imagery	4
v	A verbal factor	1
y	A factor operating in Latin and French	1

(All the positive loadings for the factors as quoted in this paper were added for these values.)

¹ FLANAGAN : *Op. cit.*, p. 25.

ALEXANDER : *Op. cit.*, p. 25.

WILSON : *Op. cit.*, p. 25.

IV.—SUMMARY AND CONCLUSIONS.

(1) An attempt has been made to find the bearing of general and special abilities upon success :

- (a) In the Scholarship Examination by which a child enters a secondary school ; and
- (b) in the School Certificate Examination taken at the end of a secondary school career.

(2) Twenty tests, some involving a high degree of the g factor, and others with various special factors, v (verbal), n (number), S_1 and S_2 (two spatial factors), and a logical factor, were formulated and analysed.

(3) The Junior Scholarship Examinations analysed involved the factors g , v , n , and x (a factor involving *attitude to work* in different proportions varying from one authority to another with the actual paper set).

(4) In the problematic type of Arithmetic Paper, the g factor was dominant, but the n factor increased in relative importance if the mechanical aspect of this subject was stressed.

(5) The g factor was dominant if Comprehension was tested in the English Paper, but if Free Composition was required, then the v factor was of greater importance.

(6) The Intelligence Tests, given in the scholarship examinations analysed, were highly saturated with g .

(7) A factor termed x which, it is suggested, involves attitude to work and persistence in character, played some part in the Arithmetic and to a less extent in the English Paper.

(8) The presence of a 'spatial' and a 'spatial imagery' factor was detected in the special tests prepared for this investigation, but these factors were not involved to any appreciable extent in the scholarship examinations analysed.

(9) The results in eleven of the subjects most frequently taken in the School Certificate Examination were also analysed. The general factor g made the greatest contribution to success, as it was operating in all the subjects analysed, excepting Art.

The x , or character factor, was next in importance since it was operating in many of the subjects analysed. The 'spatial' and 'spatial imagery' factors were respectively next and next but one in order of importance, and were operating in Art, Geometry, Geography, and Botany.

The n factor was of some importance in Arithmetic, Algebra, and to a lesser extent in Geometry.

The v factor played but a small part in the examination as a whole in that it was only found in the English Language Paper. Yet another factor, termed y , was found in Latin and French.

(10) The best method of giving practical effect to these findings must be left to future investigators, as they may be applied not only to the careers of children who gain admittance to a secondary school but also to the careers of those intending to pursue an Art, Commercial or Technical course of study.

My thanks are due to Doctor Wynn Jones for most valuable advice given throughout this investigation.

MODERN PSYCHOLOGY AND THE SPENS REPORT.

BY OLIVE A. WHEELER.

- I.—*Introduction. Terms of reference of Spens Committee. Previous (Hadow) Reports.*
- II.—*The question to be considered—the influence of educational psychology on the findings of the Spens Report.*
- III.—*The Grammar School.*
- IV.—*The Technical High School.*
- V.—*The Training of Teachers.*

I.—INTRODUCTION.

IN the critical notice of the Spens Report which appeared in the last issue of this JOURNAL, Dr. P. B. Ballard dealt with the general nature and findings of the Report of the Consultative Committee on Secondary Education. It was obviously not his intention to discuss the question which is of special interest to educational psychologists and with which I propose to deal, namely, the influence (or failure in influence) of modern psychology on the thinking and consequently on the recommendations of the Committee for the further development of Secondary Education in this country.

The Spens Report begins with two illuminating chapters, one a historical survey of Secondary (Grammar School) Education and the other a clear summary of the present provision for Secondary and Junior Technical Education in England and Wales. The facts marshalled in these chapters provide an important part of the background which obviously must be taken into account before suggestions for improvements can relevantly be made. It should be noticed, however, at the outset, that the limitation of the terms of reference of the Committee to "the organization and inter-relation of schools, *other than those administered under the Elementary Code*, which provide education for pupils beyond the age of 11+, " unduly narrowed that background and prevented the presentation of the achievements and differences of other schools for adolescents, namely, Selective Central and Senior Schools, which are at present administered under the Elementary Code. Undoubtedly the precise terms of reference, for which the Committee cannot itself be held responsible and which indeed it sometimes eluded in its deliberations, made it difficult, if not impossible, for the discoveries of modern psychology to be used freely and the problem of adolescent education to be viewed as a whole.

In the three previous Reports of the Consultative Committee on *The Education of the Adolescent* (1926), *The Primary School* (1931), and *Infant and Nursery Schools* (1933), it was clearly realized that education should be so organized as to reflect the chief stages of growth of individuals, and that the curricula and methods at each stage should be adjusted to the characteristic needs of children in that phase of development. It was also emphasized, especially in the Report on the *Education of the Adolescent*, that there should be varieties of provision to fit the chief varieties of ability and interests found at one period, for example, from 11+ to 16. It was plainly recognized that the democratic ideal of *equality* of educational opportunity was not inconsistent with the scientific differentiation of education to fit the needs of individuals, and could no longer be interpreted as implying *identity* of provision for all types of endowment. Modern psychology thus seems to have profoundly influenced the very texture of the thought of the "Hadow" Committees. Indeed the three reports together can fairly be interpreted as a big-scale, though tentative, attempt to psychologize the organization of education in this country.

II.—THE QUESTION TO BE CONSIDERED.

Is there evidence in the Spens Report of a similar recognition of the possibilities of the scientific regulation of education? Did the Committee pursue its special enquiries with an equal faith in the desirability of differentiation according to psychological needs? In particular, is there internal evidence to support the view that relevant scientific discoveries concerning adolescents freely influenced the thinking of the Committee? Or, are there signs that important psychological data have been ignored or distorted, or at least have failed to influence the findings of the Report?

The third chapter in the Report is an interesting outline of the more salient features in the physical and mental development of children between the ages of 11+ and 16+, and for this, the Committee acknowledges its indebtedness to such distinguished scientists as Dr. H. A. Harris, Professor of Anatomy in the University of Cambridge, and Dr. Cyril Burt, Professor of Psychology in the University College, London. The valuable memoranda submitted by these experts may have influenced the Committee in its judgment that the raising of the school leaving age to sixteen "must be envisaged as inevitable," and in its chief administrative recommendation that all types of schools attended by children over the age of eleven should be administered under a new Secondary Code of Regulations. There are, however, surprising omissions in the psychological approach to the study of the Committee's main problem.

For example, there is no examination of the data that has recently accumulated concerning the range and distribution of intelligence among adolescents, or the resemblances and differences between individuals in emotional experiences and educational attainments. There is no serious consideration of the development of vocational interests during adolescence. Yet in view of the Committee's recommendations concerning the amount of provision of education of the Grammar School type which should be made, and the institution of Technical High Schools, such psychological considerations would obviously have been relevant. From these omissions alone, it seems probable that the Spens Committee either had no faith in the recent discoveries of modern psychology or, in the process of its deliberations, became so snowed up with the views of experts that it failed to sift out and to make use of some of the most relevant psychological data.

The later sections of the Report confirm this general impression that, notwithstanding the inclusion of the chapter on the main features of early adolescence and of the appendices concerning Faculty Psychology and Transfer of Training, recent psychological discoveries have had a disappointingly small influence on the detailed recommendations of the Committee.

III.—THE GRAMMAR SCHOOL.

Consider, for example, the argument advanced concerning the curriculum of the Grammar School. Although there is included as an Appendix a most valuable and enlightened memorandum by Professor I. L. Kandel on the Secondary School Curriculum it is implied in the Report itself that the conception of a "liberal" or "all-round" education is necessarily associated with a discredited "faculty" psychology¹ and with nineteenth century views concerning the transfer of training. Yet few (if any) living educational psychologists have ever seriously adopted a "faculty" psychology, though there are many who would hesitate to reject the ideal of a "liberal" education during early adolescence. The very immaturity of the adolescent involves varied and partially undefined possibilities of growth and therefore makes it desirable that the Secondary School curriculum should be both flexible and well-balanced. The range and variety of interests of normal adolescents and the emotional developments characteristic of the period are sufficient grounds for the avoidance of a narrow specialism at this stage. Indeed the ideal of a "liberal" education during a period of rapid and many-sided growth needs no support from an outworn "faculty" psychology.

¹ Report: p. 70 and p. 145.

But notwithstanding its attack on the idea of a "liberal" education, the Spens Committee reaches very much the same conclusions in regard to the essential elements in the Secondary School curriculum which had previously been deduced from the consideration of the many-sided developments characteristic of the period of early adolescence. Thus in *Youth*, the irreducible minimum curriculum deduced from such psychological consideration consisted of "physical training; an art or a craft; nature study; the native language (or languages), literature, history and geography; social and religious education"¹ whereas in the Spens Report it is stated that the Secondary School curriculum should include "religious and moral teaching and training in the care of health, bodily efficiency and grace, manners and social organization" as well as "Letters: that is the use and appreciation of language, including at least some study of the native literature;" some form of Art; Handicrafts; Science (including Mathematics); History and Geography.²

The main objection to the traditional Grammar School curriculum is not that it is too "liberal" but that it is too "academic." On this issue there is no quarrel between the findings of educational psychologists and of the Consultative Committee. The Report expresses the opinion that "the studies of the ordinary Secondary Schools should be brought into closer contact than at present with the practical affairs of life."³ It regards as "the gravest defect of the present system that a boy or girl may pass through a Secondary School, having made no contact, or next to none, with one tradition—that of the arts and crafts."⁴ Indeed, it even goes so far as to affirm for the Secondary School the view put forward for the Primary School in the 1931 Hadow Report that "the curriculum should be thought of in terms of activity and experience, rather than of knowledge to be acquired and facts to be stored."⁵

On reading these expressions of the Committee's opinions, the hopes of educational psychologists naturally begin to rise. This clear realization of the unreality and rigidity of the traditional Secondary School curriculum will surely lead the Committee to tackle with all its force and ingenuity the main cause of our continued failure to adjust Grammar School education to the needs of adolescents, namely, the prevailing Secondary School examination system. The views of experts concerning the reliability (or unreliability) of examinations and their adequacy (or inadequacy) as measures of "activity and experience" and of emotional, social and moral developments, and the data accumulated by doctors and psychologists regarding the effects of examinations on the physical, mental and

¹ O. A. WHEELER: *Youth*. University of London Press, 1929.

² Report: p. 157. ³ p. 162. ⁴ p. 158. ⁵ p. 152.

moral health of Secondary School pupils will surely need to be thoroughly examined. Apparently no such investigation is attempted by the Committee, but instead in Chapter 7 there are half-hearted suggestions for slight modifications in the regulations for the First School Leaving Certificate.

The Report recommends that in future, in order to obtain a Certificate, a candidate should pass in English and in five other subjects, one of which must be a foreign language (Group II) or a scientific subject (Group III). Thus, while English is compulsory, and a foreign language and a science are semi-compulsory, the non-academic subjects (Group IV), the importance of which the Committee had previously stressed, retain their inferior examination status. To demand that a pass be achieved in each of the Groups I, II, III and IV, or to allow that the five subjects other than English should be freely chosen from any of the groups, would be a feasible position. But why should proficiency in a second language compensate for deficiency in a science, and *vice versa*, when no amount of proficiency in Art, Music or Agriculture is allowed to compensate for deficiency in a foreign language and a science?

It is argued in the Report that this discrimination between academic and non-academic subjects is due solely to the difficulty of examining the latter. If, however, æsthetic and practical activities are unexaminable, the general principle repeatedly sponsored in the Report—that the School Certificate examination should follow and not determine the curriculum—becomes an academic counsel of perfection which cannot be put into practice. Indeed, if the Secondary School curriculum should be “thought of in terms of activity and experience, rather than of knowledge to be acquired and facts to be stored,” it is difficult to resist the conclusion that the sooner the system of external examinations for First School Leaving Certificates disappears from the Secondary as well as the suggested new Technical High Schools, the better it will be for the balance and flexibility of the Secondary School curriculum and for the adjustment of secondary education to the real needs of adolescents. The scientific data concerning examinations which have already accumulated are certainly such as to justify the most serious consideration of the proposal to substitute school records (with inspections and internal examinations) for First School Leaving Certificates, even if the Second or Higher School Leaving Certificate were retained as a further condition of entrance into Universities.

IV.—THE TECHNICAL HIGH SCHOOL.

In addition to proposing that there should be differentiation of the Grammar School curriculum in the fourth and subsequent years, not

only for pupils specially interested in languages and in science and mathematics, but also for pupils who desire to give increased attention to artistic and practical subjects and for "non-academic pupils who desire to do some commercial subjects," the Report deals in Chapter 8 with another kind of differentiation. It recommends that Technical High Schools should be established, with a status equal to that of existing Grammar Schools, recruiting their pupils at 11+, and organizing a five years' course in preparation for a variety of occupations, conveniently grouped together as "Engineering."

It is proposed that where possible these Technical High Schools should be housed in Technical College or Technical Institute buildings. The Committee draws attention to the advantages of this arrangement and particularly of the sharing of expensive equipment and plant by school pupils and senior students; but apparently fails to observe that its blessing of this arrangement, and especially of the sharing of staff, cuts right across the principles embodied in its three previous reports, namely, that education at each stage should be more exactly adjusted to the special needs of the period, and that the whole educational system should be "reorganized" to make this practicable.

It will be generally agreed that there is an urgent need for the further development of technical education in this country, but nevertheless this particular proposal to select recruits at 11+ for a form of education suitable to late adolescence is open to grave objection. The differentiation of schools recommended in the first Hadow Report can be psychologically justified, for it corresponds roughly to differences in endowment, interest and attainment found at that stage. Further, while Grammar, Selective Central, and Senior Schools differ in the complexity, abstraction and scope of their curricula, they agree in their main objective of providing an education suitable to their respective pupils, but liberal in character, in the sense that encouragement is given to the general and many-sided development of each individual.

The real point of the earlier attack on the conception of a "liberal" education, to which reference has already been made, now becomes apparent. Its support is needed to justify the Committee's selection of existing Junior Technical Schools, rather than of existing Selective Central Schools, for full Secondary School status. In fairness, it should be pointed out that the Committee recommends that the first two years of the Technical High School course (corresponding to the period of general education which now precedes entrance into existing Junior Technical Schools) should continue to be general or "all round" in character. But this concession to the need for a "liberal" education in early adolescence neither diminishes the difficulty, nor justifies the

early selection of pupils for a particular form of education on a vocational, instead of a psychological, basis. At 11+ a boy's powers are so untried, his interests are so undeveloped and the organization of his emotions is so subject to fundamental change that to exact a decision from him in regard to his future vocation is certainly forestalling nature and consequently is likely to prejudice his chances of free and full development.

Such considerations cannot fail to raise doubts in the minds of educational psychologists concerning the wisdom of the recommendation for setting up Technical High Schools, recruiting entrants at the same age as Grammar, Selective Central and Senior Schools. In making this proposal, has the Committee maintained untarnished its earlier view that "a school fulfils its proper purposes in so far as it fosters growth, helping every boy and girl to achieve the highest degree of individual development of which he or she is capable"?¹ In other words, are the Technical High Schools described by the Report really conceived in the interests of the children and with the sole desire to minister to their free and full development? All the internal evidence seems to point to the conclusion that the modern tendency to psychologize education was at this point in the Report overruled by other considerations, legitimate or illegitimate.

V.—THE TRAINING OF TEACHERS.

The half-heartedness of the Committee in regard to the training of teachers is also indirect evidence of its failure to appreciate the significance of the new psychological approach to the study of education. The Report recognizes theoretically that "the curriculum, to have meaning for the learners, must be adapted to the stage of development of the pupils concerned" and that "teachers must be on their guard, to see that instruction is adapted to the interests and abilities of the pupils."² It also repeatedly emphasizes the truth that "from the educational point of view, it is not the subject, but the methods of teaching that matter."³ Yet it fails to draw the obvious conclusion that, whatever may have been customary in the past, there is now a strong case for the training of intending Secondary School teachers in the understanding of children, in the consideration of the principles of education and in the study of methods of teaching.

Among education psychologists it has long been taken for granted that teachers should be trained to approach their work from the *human*, as well as the *subject* end, for to be effective they must know their pupils as well as the subjects which they are expected to teach. Indeed, as the

¹ Ibid pp. 151-2. ² p. 78. ³ p. 173.

emphasis has gradually shifted from *subject* to *method*, as the curriculum has come to be thought of in terms of "activity and experience, rather than of knowledge to be acquired and facts to be stored," and as schools have increasingly become real societies as well as places of instruction, the need for the training of intending teachers in the principles and practice of education has been more clearly and more widely recognized.

Notwithstanding the weight of tradition in the opposite direction, this progressive movement for the training of teachers has been evident in secondary, as well as in elementary, education. The recent policy of the Board of Education in regard to University Training has encouraged Secondary as well as Elementary School teachers to spend a complete session after graduation on the professional (and human) approach to the work of teaching. Thus, according to statistics in the Report, the proportion of trained teachers at work in Secondary Schools has steadily increased from 45·7 per cent in 1927 to 59·1 per cent in 1937.¹

It might have been expected that under these circumstances a Committee advocating changes in curricula and improvements in teaching methods in Secondary Schools would have given unequivocal support to the idea of secondary training, however critical it might have had to be regarding the methods at present employed in such training. But actually the Spens Committee is disappointingly half-hearted. For example, it not only states that "teachers who are required to do specially advanced work in Grammar Schools may often spend their fourth university year more profitably in increasing their mastery of their special subjects than in following a course in the University Training Department," but even goes so far as to suggest "that in some modern schools also some relaxation of the present 'Code' requirements as to the necessity for specific training for teaching may be found desirable in exceptional cases in order to secure other specialist training or even practical experience."²

Even if there are Grammar Schools so organized that certain teachers do only advanced work—a very doubtful hypothesis in view of the complications of the time table—it is hardly consistent to assume that this is educationally desirable, when it has been recognized that "from the educational point of view it is not the subject, but the methods of teaching that matter."³ Even if it were possible to ensure in practice that such specialist, but untrained, teachers never strayed into the lower forms of the Grammar School, it would still be difficult to justify their complete neglect of the study of the characteristics of adolescents, and of the history, principles, and practice of education, before entering

¹ Ibid p. 103. ² p. 300. ³ p. 173.

upon the work of Secondary School teaching, for, as the Report itself shows, this can no longer be conceived as mere instruction, but includes also social and moral training and the encouragement of the full development of the pupils.

Of course a certain amount of flexibility in regard to the training, as in regard to the graduation, of Secondary School teachers is desirable : but this should not be interpreted to mean that either part of the two-fold qualification is unimportant. Indeed, since the limit of improvement in educational methods by haphazard means seems to have been practically reached, and developments in this direction and in curricula and internal organization are now coming more and more from systematic research, it seems probable that in the future further refinements of educational technique will result mainly from a closer alliance between psychology and education. These improvements can only be implemented in practice by extended provision for, and improved methods in, the training of teachers.

The problem of the training of the teacher is therefore at present a key problem in secondary, as in elementary, education : and the failure of the Spens Committee to deal with it can only be interpreted as a partial failure to realize the full significance of the modern scientific, and especially of the psychological, approach to the study of education.

EDUCATIONAL ABILITIES OF TRAINING COLLEGE STUDENTS.

By P. E. VERNON

(*Department of Psychology, University of Glasgow*).

I.—*Introduction.*

II.—*The inter-relations, and factorial analysis, of training college subjects.*

III.—*The use of intelligence tests.*

IV.—*Tests of interests and personality traits.*

V.—*Questionnaire on motives for choosing the teaching career.*

VI.—*Personality self-rating questionnaire.*

VII.—*Summary and conclusions.*

Appendix I.—*Statistical treatment and results.*

Appendix II.—*Student test norms.*

I.—INTRODUCTION.

Our knowledge of the organization or inter-relationships of educational abilities among children, and our success in predicting such abilities by means of psychological tests, have reached a fairly advanced stage, thanks to the researches of Burt, Spearman, and many others. That the science of adult abilities is at present a good deal more backward is probably due simply to the practical difficulties of obtaining measurements of sufficiently large numbers of adults on sufficiently large batteries of tests or examinations. However, at a training centre for teachers, as distinct from a university, most of the students do undertake the same courses of study. At one such centre the yearly influx of students includes some two to three hundred men and women, all of them graduates of a Scottish university. The writer was in the fortunate position of being able to study the marks of such groups over four years, and of having time during his Experimental Psychology classes to apply to most of them certain psychological tests.

The college subjects which we shall consider are as follows :

- (1) Psychology : educational psychology, mental testing, child development and experimental work.
- (2) Arithmetic : analysis of the fundamental arithmetical processes, alternative ways of presenting these to primary school-children.
- (3) Hygiene : human physiology, children's diseases, care of health, heating and ventilation.

- (4) Nature Study : biology and science for primary school teaching.
- (5) Education : history and theory of education.
- (6) Geography : local and world geography, climate, connections with economic conditions.
- (7) English : literature, grammar, and methods of teaching.
- (8) History : the technique of the historian, local and Scottish history.
- (9) Speech Training : physiology of speech, improvement in articulation and pronunciation, phonetics.
- (10) Teaching Skill : actual ability in the classroom as estimated by masters or mistresses of method.
- (11) Physical Training : teaching of balanced exercises, team games, gymnastics.

The following tests were given to groups of students :

- (1) Verbal Group Intelligence Tests. Four tests were applied at different times : Cattell's Scale IIIA ; the National Institute's Group Test 33 ; the Nelson-Denny Reading Test, consisting of vocabulary and paragraph comprehension (silent reading) tests ; and an unpublished group test of the usual type, constructed by the writer. The results of these tests were so similar that they will not be considered separately (see Sections II and III).
- (2) Non-verbal Intelligence Test : Stephenson's (unpublished) battery of perceptual or G-reference tests (see Sections II and III).
- (3) Educational Tests : some of the (American) Co-operative Achievement Tests for college students (see Section III).
- (4) Personality and interests questionnaires (see Sections IV to VI).

This article summarizes some of the results obtained under everyday working conditions. Although such results are, of course, much less definite and significant than those which might have been achieved under ideal experimental conditions, yet they appear to throw a good deal of light on problems of adult abilities.¹ The essential statistical details may be found in Appendix I.

II.—THE INTER-RELATIONS, AND FACTORIAL ANALYSIS, OF TRAINING COLLEGE SUBJECTS.

It is well known that almost all correlations between the scholastic and other abilities of children are positive, and that this positive overlapping has been explained by Spearman as due to a general ability factor, *g*, which runs through all the separate abilities. Certain groups of tests show additional or residual correlation, over and above that due to *g*, and this has been ascribed to so-called group factors. Thus tests

¹ A recent article in this *Journal* by J. Lawton (IX, pp. 133-144) covers some different aspects of the same general field. In so far as the two studies overlap, their results seem to be quite congruent.

involving manipulations of words are found to measure g +a verbal group factor, v . And many of the common performance tests measure g +a practical or spatial factor, which Alexander¹ calls F . The Stanford-Binet tests apparently involve both general ability and these two distinctive types of ability.² In the scholastic field Burt³ has discovered a prominent general factor, running through all the school subjects, together with subsidiary group factors limited to the three main "families" of subjects, namely the linguistic, the arithmetical, and the manual subjects. Burt's general factor coincides fairly closely with Spearman's g , and it seems quite likely that the linguistic and manual group factors correspond with v and F .

Now it has sometimes been suggested that human abilities differentiate, or become specialized, to such an extent during adolescence, that a significant general factor can no longer be recognized.⁴ However, in the educational field at least, our data show that there is positive overlapping of widely diverse abilities among training college students. All the inter-correlations (see Table III) are positive. Students who are above average in Psychology tend also to be above average in Teaching Skill, and those who are poor in Education tend also to be poor in Speech Training, and so on.

Though all but one of the correlations are statistically significant, they are rather small. The median coefficient is $+0.273$, and the highest $+0.529$. This is partly due to the fact that the reliability of training college marks is somewhat low, since there is seldom time for more than three to five hours of examinations in any one subject.⁵ Nevertheless the major proportion of the correlations may be accounted for by a general factor, which is almost as prominent at this level as among Burt's ten-year-old children. This general factor, however, is only to a small extent dependent upon the g which can be measured by intelligence tests. The correlations of verbal and non-verbal group tests with the factor are no higher than that of the Physical Training marks, namely $+0.33$ (see

¹ ALEXANDER, W. P.: "Intelligence, Concrete and Abstract." *Brit. Journ. Psychol. Monog. Suppl.*, No. XIX, 1935.

² Cf. ALEXANDER, *op. cit.*, and BURT, C.: "The Latest Revision of the Binet Intelligence Tests." *Eugen. Rev.*, 1939, XXX, pp. 255-260.

³ BURT, C.: *The Distribution and Relations of Educational Abilities*. (London: King, 1917). Also, "The Relations of Educational Abilities." *This Journal*, 1939, IX, pp. 45-71.

⁴ e.g., BUEHLER, C.: *From Birth to Maturity*. (London: Kegan Paul, 1935).

⁵ Only one reliability coefficient could be calculated, namely $+0.80$ for the Psychology marks, hence no corrections for attenuation could be applied. It is likely that Psychology is marked more thoroughly than any other subject, except Teaching Skill, since three lecturers take part in it. Hence the other reliability coefficients may be distinctly less than $+0.80$.

Table IV). Possibly then the factor consists chiefly of something in the nature of studiousness, or willingness to work. The subject which correlates most highly with, or gives the best indication of, the factor is Psychology, presumably because this is the most complex subject which students have to take, and so best brings out their willingness, or unwillingness, to work.¹

Now some of the College subjects inter-correlate more closely, some less closely, than they would if our general educational ability factor was solely responsible for the overlapping. By means of the partial correlation technique it is possible to eliminate the influence of the general factor, and to find what residual correlations there are (see Table III). A study of the partial coefficients shows that the subjects tend to fall into three main groups. First the scientific subjects (Psychology, Arithmetic, Hygiene, and Nature Study) tend to correlate positively with one another, and negatively with subjects in other groups. Secondly there is an Arts or humanistic set (Education, Geography, English, and History), and thirdly a set of practical subjects (Speech Training, Teaching Skill, and Physical Training). The groups are not entirely distinct, but the numerous "cross-currents," which run counter to the classification, are quite logical and comprehensible. For instance, Education overlaps rather highly with Psychology, although it is more akin to the Arts than to the other science subjects. Speech Training and Nature Study similarly show a marked positive association.

Here then there is evidence that the same group factors which emerged from investigations of children's abilities are present at the adult level, among a set of subjects which differ widely from those studied at the primary school age. Another means of classification is provided by Thurstone's technique of multiple factor analysis. This resolves the college marks, not into a general factor+small subsidiary group factors, but into three equally prominent common factors, or types of ability (see Table IV). The scientific type correlates most highly with (or contributes most to success in) the four subjects mentioned above. It plays a small part in Physical Training and English, the latter because English involves a certain amount of grammatical analysis. It is still less involved in Education and Speech Training, and not at all in the other subjects. The Arts or humanistic type of ability enters chiefly into the four subjects already mentioned, and into Psychology, but plays a small part in all other subjects except Physical Training. The practical type of ability is chiefly found in Teaching Skill and Speech

¹ If the reliability of Psychology marks is higher than that of other subjects, this would also lead to a higher saturation with the general factor. The correlation of these marks with total weighted college marks for all subjects except Teaching Skill is +0.878.

Training, and to a lesser extent in Physical Training, Nature Study, and English. The more theoretical subjects such as Psychology, Education, and Arithmetic contribute little to it at this stage, though they might of course become more important later in the teachers' careers.

A warning should be given here against regarding these factors—general, common, or group—as basic elements or faculties of the mind. The fact that such similar types of ability are found in so many different studies of tests and school subjects does indeed suggest the influence of a fundamental feature of mental structure. But it is safer to recall that factors are, primarily, abstractions from test scores and marks, which provide useful categories for the classification, and prediction, of abilities, and not to venture rashly upon their psychological interpretation.

III.—THE USE OF INTELLIGENCE TESTS.

All the intelligence tests which were applied correlate positively with all the college subjects, though the coefficients are not always statistically significant. Thus they do overlap to a small extent with the general educational ability factor. When the influence of this factor is removed, they give positive partial coefficients with the more theoretical subjects, and negative coefficients with most of the more practical subjects (see Table III). But the results of the verbal and non-verbal tests are decidedly different. The former agree about equally with the Science and Arts groups, the latter has no relation to the Arts group but quite high agreement with the Science group. This might have been expected since the former depend upon manipulations of words, the latter upon manipulations of abstract diagrams.

The tests agree more highly with Psychology than with any other subject, partly because Psychology has the characteristics both of a Science and of an Arts subject (as our factorial analyses show), partly because it is more reliably examined than the other subjects, and partly also because it is the only subject in which new-type examining is adopted for some of the work. Intelligence tests, being new-type in form, correlate better with new-type than with essay-type examination marks. By combining verbal and non-verbal intelligence tests a moderately high correlation with Psychology may be obtained. For example, the Nelson-Denny Reading test, which was the most successful of the four verbal tests, correlated $+0.409 \pm 0.035$ with Psychology, but only $+0.257 \pm 0.059$ with the non-verbal test. Hence the correlation of Psychology with the sum of Nelson-Denny and non-verbal tests works out at $+0.535 \pm 0.032$.

At the present time tests appear to be used in university or training college psychological departments mainly for pedagogical purposes. The extreme homogeneity of groups of adult students, due to the high standards of their selection, inevitably reduces the size of all correlations between tests and scholastic achievement, and so prevents us from ever obtaining very close predictions of achievement by means of tests. Nevertheless the coefficient of $+0.535$, just quoted, indicates that we might make some use of tests for practical prediction, even at the adult level. It should be possible to devise still more successful aptitude tests if we gave up the attempt to measure an *innate* general intellectual factor, or *g*. Instead of confining the test material to intellectual tasks which are presumed to be free from influences of upbringing or education, we might choose tasks more closely resembling the sort of tasks which students of psychology will be called upon to tackle. It would be a mistake to construct tests whose solution demanded an actual knowledge of psychology, for many of the students who eventually turn out to be the best psychologists might know no psychology at all when they took our aptitude tests. However, Earle has shown that a kind of compromise between the intelligence and the educational tests gives excellent prognoses at the secondary school level.¹ And many American universities obtain better predictions of achievement from tests like the Thorndike Intelligence Examination, the Nelson-Denny Reading, and some of the Co-operative tests, all of which include educational material, than from either intelligence tests or ordinary examinations.

The Co-operative Achievement Tests, which are constructed mainly for purposes of educational guidance of incoming university students, often contain too much specifically American material to be suitable for British students. However, the writer has obtained promising results with the following five tests:

- (a) General knowledge of literature.
- (b) General knowledge of art.
- (c) General scientific knowledge.
- (d) General mathematical knowledge.

These four are all parts of the *Co-operative General Culture Test*.

(e) Test of interpreting facts. This test, taken from the *Co-operative Test of Social Studies Abilities*, consists of a series of paragraphs, each followed by six statements. The student has to decide whether each statement is a legitimate deduction from the facts given in the paragraph, whether it goes beyond the facts, or whether it contradicts them.

¹ EARLE, F. M. : *Tests of Ability for Secondary School Courses*. (University of London Press, 1936.)

The tests were given to a group of some 240 students, and a shortened time limit of twenty minutes for each test was imposed. The correlations with total Psychology marks were : (a) $+0.209$, (b) $+0.190$, (c) $+0.207$, (d) $+0.155$, (e) $+0.364$. Combining the first four, and giving double weight to (a) and (c), yields a correlation of $+0.295 \pm 0.041$. This coefficient is almost as high as those given by verbal intelligence tests. Combined with Test (e), it rises to $+0.408 \pm 0.037$. Such results suggest that, with adequate research, some useful aptitude tests for predicting adult capacities in Psychology or in general educational ability might be constructed.

VI.—TESTS OF INTERESTS AND OF PERSONALITY TRAITS.

We suggested above that the general educational ability factor may consist largely of studiousness or willingness to work. Hence it is worth enquiring whether cognitive aptitude tests might be supplemented with tests of interests or temperament. Several American studies with extraversion tests or personality questionnaires have shown a slight tendency for better students to be above average in introversion. More definite results have come from the application of tests of interests or values. In collaboration with G. W. Allport, the writer gave the *Study of Values* test to a small group of professors, lecturers and research students in psychology, and found that they possessed a distinctive pattern of interests.¹ They were especially strong in theoretical or intellectual and æsthetic or cultural interests, and relatively lacking in economic or business, political or power-seeking, and religious interests. This general picture is confirmed and extended by an examination of Strong's Vocational Interest Blank.² Strong's test consists of some four hundred varied items, each of which is checked by the subject according as he likes it, dislikes it, or is indifferent to it. The scoring keys show that professional psychologists are chiefly characterized by the following items. They would like to be authors, poets or sculptors, university professors, mathematicians, scientific research workers, doctors or school-teachers, but would hate to be salesmen or business agents of any description. Among school subjects they most enjoyed the biological ones. They are interested in Socialism, are tolerant of (indifferent to) most types of "peculiar" people, but dislike self-made men and born leaders. Also they dislike organizing or contributing to charities. They

¹ VERNON, P. E., and ALLPORT, G. W. : "A Test for Personal Values." *Journ. Abnormal and Social Psychol.*, 1931, XXVI, pp. 231-248.

² STRONG, E. K. : *Vocational Interest Blank for Men*. (Rev. edit., 1938. Stanford University Press.)

much prefer developing plans to executing plans ; for instance they would rather work out the theory of a machine than actually make the machine or try to sell it. Again they are happier if given charge of technical operations than in administrative posts. They value freedom to work out their own ideas rather than advice or direction from superiors.

These, of course, are average tendencies which admit of numerous exceptions ; and they are characteristic of American, not necessarily of British, psychologists. Further, we must not assume without proof that more able students have interests which resemble those of professional psychologists. Nevertheless the same type of test could certainly be worked out in this country, and we might expect to find that the better students would tend to give a different pattern of likes and dislikes from the poorer students.

Several minor experiments on the interests, attitudes and personality traits of training college students were carried out by the writer. Thouless's questionnaire on certainty of religious beliefs¹, and Eaglesham's list of educational aims or ideals², yielded results similar to those of their authors, but could not be found to have any noteworthy connections with student abilities. A survey of political opinions indicated that these range as widely among students as among the general population. There was no evidence that future teachers as a class are either revolutionary Socialists or hidebound Tories. Again, these opinions seemed irrelevant to ability. Some of the results of Valentine's questionnaire on reasons for choosing the teaching profession, and of Boyd's personality questionnaire, are, however, of sufficient significance to be reported here.

VII.—QUESTIONNAIRE ON MOTIVES FOR CHOOSING THE TEACHING CAREER.

Valentine's questionnaire has been fully described in this *Journal*.³ The adaptation which was made by the writer differed in only two important respects, first in the substitution of ratings for rankings of the various items, and secondly in the omission of the item—"Board of Education or other grant giving chance of university education," which is very influential among English teachers, but plays no part in Scotland. The adapted questionnaire was as follows. It was explained to, and discussed with, the students before they answered, and they were assured that their answers would be anonymous.

¹ THOULESS, R. H. : "The Tendency to Certainty in Religious Belief." *Brit. Journ. Psychol.*, 1935, XXVI, pp. 16-31.

² EAGLESHAM, E. J. R. : "An Enquiry concerning the Practicability of Typical Educational Aims." This *Journal*, 1937, VII, pp. 23-38.

³ VALENTINE, C. W. : "Reasons for the Choice of the Teaching Profession by University Students." This *Journal*, 1934, IV, pp. 237-259.

QUESTIONNAIRE ON MOTIVES FOR TEACHING.

What were the reasons which originally led to your taking up the career of teaching, and what are the reasons that chiefly appeal to you now?

In *each* column, below, write a "3" after those motives which had (or now have) a major influence. Write a "2" after those which were (or are) the next most important. Write a "1" after any reasons which had (or now have) a minor influence, and a "0" after those which did not (or do not) affect you at all.

It is suggested that you should write between one to three "3's"; one to three "2's"; two to six "1's"; and four to twelve "0's."

	<i>Original Motive.</i>	<i>Present Motive.</i>
(a) Economic desirability of the profession, including security and pensions, as well as salaries
(b) Social status of the profession
(c) The long holidays
(d) Ideals: e.g., possibility of influencing young people, or interest in education
(e) Feeling of special fitness or capacity for the work of teaching
(f) Enjoyment of feeling of leadership or power
(g) Liking for teaching, more general or vague than (d), (e) or (f)
(h) Fondness for children
(i) Desire for a profession connected with some favourite subject of study
(j) Influence of parents or other relatives
(k) Influence of your own friends at school or university
(l) Deliberate influence or persuasion of one of your own teachers
(m) Influence of a teacher (unknown to him or her) who was greatly admired
(n) Stop-gap: while looking round for something else, or waiting for some stronger interest to develop
(o) No prospect of any alternative career
(p) Inability to change now. (Answer in second column only)
(q) Any other not covered by the above. (Write below.)

The average ratings of 120 men and 116 women students are given in the first two columns of Table I, and their combined results in the third column. For comparison Valentine's results (which were based on

rankings by 195 men and 153 women in English training centres), have been converted arithmetically into a scale with a similar range and average, and are listed in the fourth column. In each column are given the ratings both of original and of present motives.

TABLE I.
AVERAGE RATINGS OF MOTIVES FOR CHOOSING TEACHING CAREER.

MOTIVES.	Men.		Women.		Combined Scottish.		Combined English.	
	Orig.	Pres.	Orig.	Pres.	Orig.	Pres.	Orig.	Pres.
(a) Economic desirability ..	2.1	2.2	1.7	1.9	1.9	2.0	1.7	1.3
(b) Social status	0.8	0.6	0.6	0.5	0.7	0.6	0.8	0.5
(c) Long holidays	1.4	1.4	1.4	1.4	1.4	1.4	1.1	0.8
Board of Education grant for university training ..	—	—	—	—	—	—	1.5	0.4
(d) Educational ideals ..	1.3	2.1	1.4	2.3	1.4	2.2	0.9	1.9
Interest in education ..	—	—	—	—	—	—	0.6	1.5
(e) Feeling of fitness ..	1.1	1.5	1.1	1.1	1.1	1.3	1.0	1.2
(f) Enjoyment of power ..	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.6
(g) Liking for teaching ..	1.4	1.7	1.5	1.7	1.5	1.7	1.9	1.9
(h) Fondness for children ..	1.0	1.5	1.8	2.2	1.4	1.8	1.2	1.6
(i) Interest in some favourite subject of study ..	1.6	1.8	1.5	1.5	1.5	1.6	2.0	2.1
(j) Parents' influence ..	1.2	0.7	1.3	1.0	1.2	0.8	1.2	0.3
(k) Friends' influence ..	0.3	0.2	0.4	0.3	0.3	0.2	—	—
(l) Persuasion by teacher ..	0.2	0.1	0.3	0.1	0.2	0.1	0.6	0.1
(m) Influence of admired teacher	0.6	0.4	0.4	0.3	0.5	0.3	0.9	0.2
(n) Stop-gap	0.7	0.4	0.6	0.6	0.6	0.5	0.3	0.2
(o) No alternative prospect	0.5	0.6	0.4	0.6	0.5	0.6	0.6	0.7
(p) Inability to change now	—	0.4	—	0.6	—	0.5	—	0.7

Certain differences between the Scottish and English figures arise from the differences in the questionnaires. In all other respects the results are so similar that almost the whole of Valentine's conclusions apply here also. In two ways, however, the study was carried a stage further than was Valentine's. First, each student's ratings on original motives *a—c* and *j—o* were subtracted from his ratings of motives *d—i*, so yielding a rough measure of the adequacy and desirability of his motives. These scores ranged from +12 to -13. The students were then classified according as they had answered 0, 1, 2 or 3 to (*p*) "Inability to change now," i.e., according to their present satisfaction or dissatisfaction with their career. The average "desirable motives" score for each of these four groups are given in Table II.

TABLE II.

COMPARISON OF DESIRABILITY OF MOTIVES FOR ORIGINALLY CHOOSING THE TEACHING CAREER WITH PRESENT DISSATISFACTION.

<i>Score on motive (p) (present dissatisfaction).</i>	N.	<i>Average score on original desirable motives.</i>	P.E. _m
0	169	+0.91	±0.23
1	30	-0.93	±0.62
2	23	-3.31	±0.61
3	14	-5.43	±0.81

A definite association is seen between undesirable motives for the original choice of the career and present dissatisfaction. However, the variability is large, and the differences between the average scores of the 0's, 1's, 2's and 3's are only about two and a half times their P.E.s. The difference between the 0's and 2's, and the difference between the 1's and 3's, are fully significant statistically.

Secondly, most students had recorded on their questionnaires the approximate grades they had obtained in a psychology examination and in an intelligence test (Cattell's Scale IIIA). The ratings of each motive were compared with these grades, and a slight negative agreement appeared with (a), (c), (n), (o), and (p), that is with the economic and the "no alternative" types of motive. The ratings on these categories, including both original and present motives, were therefore summed and correlated with grades, yielding a coefficient of -0.197 ± 0.051 . with intelligence test scores the correlation was -0.096 ± 0.058 . The connection is therefore very small, but it indicates that undesirable motives for entering the teaching career do have something to do with poor work during the course of training.

VIII.—PERSONALITY SELF-RATING QUESTIONNAIRE.

Boyd's Personality Questionnaire¹ includes 120 questions which have to be answered either "Yes," "Yes?", "0 (Doubtful)," "No?", or "No." The items are then grouped by the scorer under twenty general headings or tendencies, six items falling under each heading. But they are so arranged in the questionnaire that the subject is not likely to guess

¹ The questionnaire is not yet published, but a brief account of it may be found in the present writer's *Industr. Health Res. Board Rep.* No. 83, together with a factor analysis of the tendencies which it aims to measure.

what tendencies they refer to. Examples of these tendencies are as follows :

<i>Tendency.</i>	<i>Sample Question.</i>
(1) Tenseness	Do you find it hard to rest quietly and do nothing whatsoever ?
(2) Pressure needed for action (undependable).	Do you always have a rush at the last minute ?
(3) Depression, melancholia ..	Do you ever feel that life is rather a burden ?
(4) Lack of ability to concentrate.	Does your mind wander badly when you are reading something difficult ?

The five possible answers are marked 4, 3, 2, 1 and 0, respectively, so that the subject's scores on each tendency range from 24 (six "Yes" answers) to 0.

The obvious objection to a test of this type, which the writer has discussed at length elsewhere,¹ is that subjects are likely to be suspicious and to conceal their real feelings and motives, if indeed they really know them themselves. To reduce the weight of this objection as far as possible, the questionnaire was answered anonymously by 156 of the writer's students. Each student also recorded the grades which he had received in a psychology examination and in an intelligence test.

On several of the tendencies differences were found between good and poor students. In Tenseness (No 1 *supra*) the better students' scores were distinctly higher, on the other three tendencies cited they were generally lower. Other tendencies suggested that they are liable to "inferiority complexes" and phobias, to be introverted rather than extraverted. Yet they are fairly stable emotionally and have good control over their moods. This picture fits in well with the intellectuality revealed by the Strong Interest Blank.

Scores on the four tendencies listed above were taken, those on Nos. 2, 3, and 4 were reversed, and the totals were correlated with Psychology marks. The resulting coefficient of $+0.443 \pm 0.042$ shows a considerable overlap between ability and personality self-ratings. Further, this overlap is largely independent of that between ability and aptitude tests, for the correlation between the same self-ratings and Cattell Scale IIIA scores was only $+0.166$. Instead of using the scores on a few of Boyd's general tendencies, it would be better to carry out an item-analysis of the whole questionnaire, and thereafter to score those items which were found empirically to correlate best with ability.

¹ *loc. cit.*

Whether a questionnaire which was not answered anonymously would be equally successful, it is difficult to say. If the significant questions were mixed up with other, non-significant ones, their aim would not be very obvious, and students would not be likely to realize that their studiousness was being tested, nor answer accordingly. But only further experiment can show whether these suppositions are justified.

IX.—SUMMARY AND CONCLUSIONS.

The marks of large groups of adult students on a variety of training college subjects were studied, and were compared with their scores on certain group tests of intelligence and of personality traits. The following main conclusions emerged :

(1) Adult educational abilities are in general inter-related in a similar manner to those of children. Thus all inter-correlations are positive.

(2) The application of factorial analysis shows the existence of three main types of ability, namely scientific-mathematical, Arts or humanistic, and practical.

(3) The general educational ability factor only correlates to a small extent with the *g* measured by intelligence tests, either of the verbal or the non-verbal type. Probably temperament, interests and work attitudes play a large part in it.

(4) Verbal intelligence tests correlate about equally with Science and with Arts subjects. Non-verbal tests correlate much better with Science than with Arts subjects. Both give very low correlations with practical subjects.

(5) Moderately good predictions of achievement in psychology may be obtained by combining the results of verbal and non-verbal tests. Educational aptitude tests, such as are widely used in American universities, also give promising results.

(6) Psychologists are found to possess a distinctive pattern of interests. This might form the basis of a fresh approach to the prediction of aptitude for psychology.

(7) Students who admit, in an anonymous questionnaire, to having entered the teaching profession for economic or other inadequate reasons show a slight tendency to poorer work in psychology.

(8) Those who chose the teaching career for undesirable reasons tend to be more dissatisfied with their choice now.

(9) Answers to a personality questionnaire show the better students to be more tense, more dependable, better at concentrating, but also less liable to depression and emotional instability than poorer students. It would seem then that both the cognitive and the orectic factors which contribute to ability in psychology may be analyzed and measured with some degree of success.

APPENDIX I.—STATISTICAL TREATMENT AND RESULTS.

The raw correlations between college subjects and intelligence tests, shown in Table III, are weighted averages of the coefficients among three groups of students. The average size of the population is 560. P.E.'s run from $\pm .022$ to $\pm .035$, except in the case of the Non-verbal Intelligence Test, which was taken by only 125 students. The P.E.'s of its coefficients run from $\pm .060$ to $\pm .071$.

TABLE III.

INTER-CORRELATIONS OF TRAINING COLLEGE MARKS AND INTELLIGENCE TESTS.

Original coefficients above the diagonal; partial coefficients (with general educational ability factor held constant) below the diagonal.

	Psychology	Arithmetic	Hygiene	Nature Study	Education	Geography	English	History	Speech Training	Teaching Skill	Physical Training	Verbal Intel.	Non-verbal Intel.
Psychology	—	+ .46	+ .53	+ .42	+ .50	+ .33	+ .44	+ .30	+ .32	+ .29	+ .21	+ .34	+ .36
Arithmetic	+ .21	—	+ .38	+ .32	+ .16	+ .24	+ .27	+ .12	+ .14	+ .17	+ .17	+ .20	+ .33
Hygiene	+ .18	+ .17	—	+ .39	+ .28	+ .24	+ .32	+ .13	+ .24	+ .29	+ .23	+ .18	+ .19
Nature Study	— .07	+ .07	+ .06	—	+ .28	.31	+ .33	+ .27	+ .41	+ .27	+ .18	+ .20	+ .33
Education	+ .20	— .09	— .02	— .06	—	+ .27	+ .38	+ .25	+ .27	+ .25	+ .13	+ .25	+ .15
Geography	— .08	+ .02	— .06	+ .01	+ .02	—	+ .33	+ .28	+ .23	+ .27	+ .11	+ .15	+ .04
English	— .04	— .01	— .05	— .07	+ .09	+ .05	—	+ .24	+ .37	+ .32	+ .22	+ .22	+ .18
History	+ .01	— .07	— .13	+ .05	+ .06	+ .11	— .00	—	+ .24	+ .30	+ .02	+ .15	+ .02
Speech Training	— .17	— .15	— .11	+ .12	— .03	— .05	+ .04	+ .03	—	+ .43	+ .26	+ .09	+ .23
Teaching Skill	— .22	— .10	— .03	— .10	— .05	+ .01	+ .02	+ .11	+ .18	—	+ .29	+ .10	+ .05
Physical Training	— .06	+ .03	+ .05	— .02	— .05	— .06	+ .02	— .13	+ .10	+ .14	—	— .09	+ .12
Verbal Group Intelligence Tests	+ .14	+ .06	— .01	— .01	+ .09	— .01	+ .02	+ .02	— .13	— .11	— .02	—	+ .36
Non-verbal Intelligence Test	+ .18	+ .22	— .00	+ .17	— .02	— .15	— .03	— .13	+ .02	— .16	+ .02	+ .28	—

Saturations of the college subjects with the general educational ability factor were calculated by Spearman's method (*Abilities of Man*, Appendix, p. 16), and are listed in the first column of Table IV. Partial inter-correlations with this factor held constant are given below the diagonal in Table III¹. Those which amount to three or more times their P.E.'s are in heavy type.

TABLE IV.

RESULTS OF FACTOR ANALYSES OF COLLEGE SUBJECTS AND INTELLIGENCE TESTS.

	Saturation with General Educational Ability Factor.	Partial Correlations with the sums of the three main groups of subjects.			Factor Loadings obtained by Thurstone's Method.			Factor Loadings after rotation of axes.		
		Scientific Group.	Humanistic Group.	Practical Group.	General Factor.	Theoretical vs. Practical.	Materialistic vs. Humanistic.	Science.	Artis.	Practical.
Psychology	+ .75	+ .17	+ .05	— .22	+ .77	+ .17	— .15	+ .61	+ .49	+ .24
Arithmetic	+ .44	+ .25	— .07	— .11	+ .51	+ .35	+ .07	+ .57	+ .12	+ .13
Hygiene	+ .57	+ .22	— .12	— .05	+ .57	+ .15	+ .07	+ .44	+ .22	+ .27
Nature Study	+ .61	+ .03	— .03	— .00	+ .62	+ .03	+ .12	+ .39	+ .23	+ .46
Education	+ .52	+ .01	+ .10	— .06	+ .54	— .10	— .33	+ .12	+ .57	+ .15
Geography	+ .49	— .04	+ .10	— .05	+ .46	— .15	— .15	+ .08	+ .40	+ .25
English	+ .62	— .08	+ .07	+ .02	+ .60	— .15	— .08	+ .18	+ .42	+ .36
History	+ .40	— .06	+ .09	+ .01	+ .39	— .20	— .19	— .01	+ .42	+ .21
Speech Training	+ .55	— .14	+ .00	+ .19	+ .55	— .29	+ .25	+ .13	+ .19	+ .82
Teaching Skill	+ .54	— .20	+ .03	+ .22	+ .52	— .07	+ .20	+ .01	+ .25	+ .64
Physical Training	+ .39	+ .00	— .10	+ .16	+ .39	— .07	+ .28	+ .20	— .01	+ .39
Verbal Intelligence	+ .93	+ .08	+ .06	— .13	+ .39	+ .20	— .17	+ .32	+ .29	— .01
Non-verbal Intelligence	+ .33	+ .25	— .16	— .06	+ .41	+ .46	+ .07	+ .60	+ .04	+ .00

¹ Thoulens's objections to the use of partial correlation were disregarded, because the procedure adopted here simply follows that of Burt, and because they appear to invalidate all factorial studies. Cf. Thoulens, R. H.: "The Effects of Errors of Measurement on Correlation Coefficients." *Brit. Journ. Psychol.*, 1939, xxix, pp. 383-403.

There is still much controversy as to which of the many factorial methods may best be used for the resolution of sets of scores or marks into their underlying components. But the following statements provide, it is hoped, a fair summary of the present position :

(1) The Two-factor theory, according to which one general factor and specifics will account for all correlations, is only applicable when either the population is so small that residual correlations and tetrad differences are unreliable, or the tests are carefully selected (as in Brown and Stephenson's investigation¹) to avoid residual overlap. Under these conditions Spearman's original technique is the most accurate. However, neither condition holds in the present study, and the correlations do not approximate to hierarchical structure.

(2) In studying the relations between interests or temperamental traits, significant negative correlations are often found. Here, then, a general factor cannot readily be extracted, and its psychological meaning would be obscure.² More appropriate is a scheme of multiple common factors, each entering into several, but none into all, of the tests. Thurstone's or Hotelling's techniques, or Burt's method of least squares, should be used.

(3) In almost all other instances tests may be analyzed either into multiple common factors, or into a prominent general factor, smaller group factors and specifics. These different techniques may lead to results which are superficially contradictory, but which are, as Burt has shown, entirely logically consistent.³ Several recent investigations have indicated that alternative analyses may legitimately be applied to the same tests, and our own data will provide further confirmation.⁴

A group-factor technique such as Burt's or Holzinger's,⁶ requires that the group factors should each be sharply restricted to a few tests, and each test should be loaded with the general factor and one group factor only. In the present instance the students' college marks show too many "cross-currents," as mentioned above, for either of these techniques to be fully effective. The writer has, therefore, adopted the more primitive method of calculating from the partial correlations, the correlation of each subject with the three main groups of subjects, using Spearman's formula for correlation of sums. The results are given in Columns 2-4 of Table IV. This is not a table of group factor saturations, since it includes both positive and negative loadings, all of them low, instead of fairly high positive and zero loadings. But it serves the same purpose as a true group factor analysis.

Thurstone's multiple factor method was next applied to the original correlation matrix. Three factors were found to cover almost the whole of the communalities, the only important residual correlation being between the two intelligence tests. The analysis was revised twice until the guessed communalities approximated to the obtained communalities, and the resultant factor loadings are shown in Columns 5-7 of Table IV.

It will be seen that the first factor is almost identical with the general factor obtained by Spearman's technique (Column 1), in spite of the fact that the intelligence tests were not included in determining the general factor, but were included in the Thurstone analysis. The average deviation of the loadings is $\pm .025$. The second, bi-polar, factor apparently represents a theoretical *vs.* practical dimension, and the third might be termed a materialistic *vs.* humanistic one. Thurstone has shown that more meaningful factors, whose loadings are all positive or zero, can be obtained by rotation of axes. The application of this operation gave the factors listed in the last three columns of Table IV.

¹ BROWN, W., and STEPHENSON, W.: "A Test of the Theory of Two Factors." *Brit. Journ. Psycho.* 1933, xxiii, pp. 352-370.

² Cf. VERNON, P. E.: "The Assessment of Psychological Qualities by Verbal Methods." *Industr. Health Res. Board Rep.*, No. 83, 1938.

³ BURR, C.: "Factor Analysis by Sub-matrices." *J. Psychol.*, 1938, vi, pp. 339-375. Also, "The Relations of Educational Abilities." *This Journal*, 1939, ix, pp. 45-71.

⁴ Cf. Spearman's re-working of Thurstone's primary ability data, and Davies's study of Stephenson's data. SPEARMAN, C.: "Thurstone's Work Re-worked." *Journ. Educ.* 1937, xxi, pp. 1-12. Also, "The General Factor in Correlations Between Persons." *Brit.*

Note that these three factors account for 11.5 per cent, 13.5 per cent and 11.8 per cent of the total variance, whereas the group-factor technique ascribed 27 per cent of the variance to the general factor. Nevertheless the relative loadings of the multiple factors are very similar to those obtained by the rough group-factor technique. Indeed the correlation between the coefficients in Columns 2-4 and Columns 8-10 is $+0.862 \pm .028$. It would seem to the writer that there is little to choose between the alternative analyses from the standpoint of practical usefulness of results, though the general+group factor scheme is perhaps more in accord with commonsense notions of abilities.

APPENDIX II.—STUDENT TEST NORMS.

The following tables of scores may be useful to lecturers in training centres and universities, who wish to apply the tests. In each instance the highest and lowest scores (the 100th and 0th percentiles), and the 99th, 90th, 75th, 50th, 25th, 10th, and 1st percentiles are listed. The students were, so far as is known, a representative sample of Scottish university graduates, the majority being aged twenty-one or twenty-two years. Roughly one quarter had honours, and three quarters ordinary degrees. Separate results for men and women are quoted only when there was a significant difference between their average scores. For a few tests figures are given also for non-graduate women students, mostly aged eighteen or nineteen years.

TABLE V.
STUDENT TEST NORMS.

m=men graduates; w=women graduates; ng=women non-graduates.

Test.	Students.	Percentiles.									
		0	1	10	25	50	75	90	99	100	
(1) CATTELL INTELLIGENCE SCALE IIIA	{ 139m 120w }	73	76	87	98	101	108	115	126	134	
(2) STEPHENSON NON-VERBAL G TEST :											
(a)	{ 68m 60w }	13	19	25	30	38	47	53	63	64	
(b)	{ 54m 64w }	17	20	32	37	43	50	54	60	66	
(c)	{ 60m 60ng }	14	14	25	31	39	46	56	60	61	
(3) N.I.I.P. GROUP TEST 33 :											
(a)	{ 75m 65w 58n 60w }	59½	86	128½	145½	158	165½	174	179	185½	
(b)	{ 58m 60w }	124	134½	148	158½	166	174	178	182	183	
(c)	{ 118m 118ng }	98	110½	124½	131	143	153½	164	172	173	
(4) NELSON-DENNY READING TEST A	{ 137m 127w }	62	68	87	103	114	129	140	158	160	
(5) CO-OPERATIVE GENERAL CULTURE TEST, FORM O :											
II.—Foreign Literature ..	{ 144m 96w }	22	26	32	42	64	89	109	124	131	
		18	20	31	40	55	67	80	110	131	
III.—Art	{ 129m 96w }	28	32	41	51	68	84	99	120	123	
		11	20	36	45	58	70	85	107	114	
IV.—Science	{ 146m 96w }	10	13	28	32	42	54	67	83	85	
		16	17	23	27	38	45	52	68	69	
V.—Mathematics	{ 143m 94w }	5	13	20	27	36	46	52	56	59	
		9	11	16	20	27	34	41	48	48	
(6) SCHONELL SPELLING .. .	{ 68m 63w }	37	38	44	49	52	55	57	58	58	
(7) McADORY ART JUDGMENT (half the test)	{ 72m 113w and ng }	55	59	66	69	75	81	93	101	106	
		50	53	68	75	81	89	95	107	114	

(1) *Cattell Intelligence Test, Scale IIIa*.—Note that the median score of 101 corresponds to an I.Q. of 146 according to Cattell's norms. This seems unduly high, until it is remembered that Cattell's Test has a S.D. for I.Q.'s of 25. Converted to an I.Q. scale with the more usual S.D. of 15 or 16, the median becomes the quite reasonable figure of 128 to 130.

(2) *Stephenson's Non-verbal G Test*.—A simplification was introduced into the scoring of the third sub-test of this battery. Each correct item was assigned 3 instead of 6 marks. This reduces the maximum possible score for the whole battery to 93.

For this, and for the following tests, separate figures are available: (a) for "naive" graduate students who had scarcely any experience of tests; (b) for "sophisticated" graduate students who had had a term's course on mental testing; (c) for "naive" non-graduate women students (cf. the writer's investigation of test-sophistication¹).

(3) *National Institute of Industrial Psychology, Group Test 33*.—The same three sets of figures are available. Note that here the graduate students (a) score more highly than non-graduate (c), whereas there was no such difference on the previous test. This suggests that the two groups may be of equal *g*, but that verbal tests are markedly affected by university training.

(4) *Nelson-Denny Reading Test, Form A*.—The median for American fourth-year students is 96, that is far lower than the Scottish figure.

(5) *Co-operative General Culture Test, Form 0, 1938*.—Part I (Social Studies) was omitted. The remaining parts are: (II) Foreign Literature, (III) Art, (IV) General science, (V) Mathematics. For each of these parts twenty minutes was allowed, instead of thirty—the American time limit. Nevertheless the American medians for second-year students are decidedly lower on all the tests but Science, namely 39, 43, 53, and 18. Compared with American standards, the scientific education of Scottish students is rather backward. Note that on every test, not only on the science and mathematics parts, women are much inferior to men in their general knowledge. This result, and others given in Section III, above, are comparable with those recently published by Inman.²

TABLE VI.
STUDENT TEST NORMS FOR KWALWASSER-DYKEMA MUSIC TESTS.
m=men, and w=women, graduate students.

Test.	Students.	Percentiles.									American school median.	Average for 22 musicians.
		0	1	10	25	50	75	90	99	100		
I.—Tonal Memory ..	61m	12	12	13	16	18	21	22	24	25	15	22.9
	41w	12	13	16	18	19	22	23	24	25		
II.—Quality Discrimination	69m	12	13	20	22	23	25	26	29	30	21	26.4
	56w	12	13	20	22	23	25	26	29	30		
III.—Intensity Discrimination ..	69m	19	20	22	23	24	25	26	28	28	22	23.9
	56w	19	20	22	23	24	25	26	28	28		
IV.—Tonal Movement ..	69m	6	7	10	15	21	26	28	29	30	16	26.1
	56w	13	14	14	18	22	26	28	29	30		
V.—Time Discrimination	61m	10	15	17	20	21	22	23	24	25	17	21.4
	41w	10	15	17	20	21	22	23	24	25		
VI.—Rhythm Discrimination	69m	12	13	16	18	19	21	22	23	23	17	20.4
	56w	12	13	16	18	19	21	22	23	23		
VII.—Pitch Discrimination	61m	19	19	22	25	27	29	32	34	35	26	26.6
	41w	19	19	22	25	27	29	32	34	35		
VIII.—Melodic Taste (Single test) ..	104m	4	5	6	7	8	9	9	10	10	—	6.7
	71w	4	5	6	7	8	9	9	10	10		
(Double test) ..	61m	9	10	12	14	15½	17	18	20	20	13	—
	41w	9	10	12	14	15½	17	18	20	20		
IX.—Pitch Imagery ..	69m	8	10	13	15	16	19	21	24	25	14	24.0
	56w	11	12	14	16	18	21	22	23	24		
X.—Rhythm Imagery	61m	10	12	17	18	19	21	22	24	24	17	22.3
	41w	15	16	18	19	20	22	22	23	24		

¹ VERNON, P. E.; "Intelligence-test Sophistication." *This Journal*, 1936, viii, pp. 237-244.

² INMAN, D. M.; "General Knowledge and Intelligence." *This Journal*, 1939, ix, pp. 72-87.

(6) *Schonell Adult Spelling Test*.¹—Schonell's averages for English students and educated adults were 43 to 45. Scottish school standards on the three R's are higher than English, and seem to remain so at the adult level. If "synchronize" may be spelt "synchronise" all the figures in the table should be raised by 1.

(7) *McAdory Art Judgment Test*.—The test contains seventy-two items. Different groups of students took only the first, or the last thirty-six, but their results were so similar that they have been combined. Non-graduate women were found to score as highly as graduates and are, therefore, included in their norms. The American median for college students is claimed as 90, and the writer obtained a figure almost as high as Harvard University. It is difficult to know why British students should be so inferior.

(8) *Kwalwasser-Dykema Music Tests*.—These are a much more interesting set of gramophone records than the Seashore tests, though less reliable.² The American medians for secondary schools pupils are much lower than our results, and are included in Table VI. Results are also given for twenty-two musicians—students and staff. Note that these are scarcely any higher than the results for students in general on Tests III, V, VI, and VII, that is on the tests for the more elementary auditory capacities, but distinctly higher on the tests involving more musical material. Sex differences, and differences between American (school) and Scottish (adult) medians are also much smaller on the former tests than the latter.

Test IX (Pitch Imagery) was converted into a sol-fa test for students who did not know staff notation. Their results were almost identical with the results of those who took the staff version, hence no separate norms are quoted.

¹ SCHONELL, F. J.: "Ability and Disability in Spelling Amongst Educated Adults." *This Journal*, 1936, vi, pp. 123-146.

² FARNSWORTH, P. R.: "Further Studies of the Seashore and Kwalwasser Batteries." *Genl. Psycho. Monog.*, 1934, xv, 1-94.

SEX DIFFERENCES AMONG INDIAN CHILDREN IN THE BINET SIMON TESTS.

By V. V. KAMAT

(Department of Education, University of Edinburgh, Moray House,
Edinburgh).

- I.—Introduction.
- II.—Comparison of the mental ages of boys and girls at various ages.
- III.—Comparison of histograms of boys' and girls' I.Q.s.
- IV.—Comparison of boys' and girls' performances in individual tests and in predominant abilities.
- V.—Summary and conclusions.

I.—INTRODUCTION.

WHILE standardising the Binet Simon Tests for Indian Children from the Stanford version, certain sex differences were noticed and these are given in brief in the present paper. It is interesting to note how these tests work in different societies and environments. Although the Indian location of the tests in general agrees remarkably well with that of Binet, Terman, or Burt¹, certain significant sex differences were noticed, which it is of interest to compare with the observations of others. To begin with it must be remembered that in Indian society the girls grow up in a different environment from that of boys. They spend more of their time in the company of their mothers and other female members of the family, the majority of whom are still illiterate, whereas the boys spend more of their time with their fathers who are more literate. Further, it is still the fashion to bring up girls to be good mothers and housewives and to bring up boys for a literary career. All the boys and girls in this study, however, except those of pre-school ages, were school-going children.

¹ See this *Journal*, Vol. IV, Part III, November, 1934, pp. 296-309. For a complete description of the experiment see the writer's *Measuring Intelligence of Indian Children* (Oxford University Press).

The agreement between the order of difficulty of the tests in the four revisions alluded to was calculated by using Spearman's rank-differences formula with squared differences. Only tests which were very nearly alike were taken into consideration. As the Stanford revision was more closely followed than the others the number of tests considered from this revision was the largest. The coefficient of correlation thus calculated between present revision and Binet (1911) was .966 (39 tests); present revision and Terman's Stanford revision .984 (67 tests); and present revision and Burt's London revision .970 (43 tests). It should be remembered, however, that these values are no measure of the agreement of placing of the tests in the different age groups, since it is possible to obtain a perfect correlation of 1, although the placing of tests may vary very much. Thus keeping the order of the tests the same we may put the tests of the third year in the fourth, those of the fourth in the fifth, and so on, and still obtain a correlation of 1.

252 *Sex Differences among Indian Children in Binet Tests*

II.—COMPARISON OF THE MENTAL AGES OF BOYS AND GIRLS AT VARIOUS AGES.

The average mental ages of boys and girls for the several chronological ages are given below :

<i>Chronological Ages.</i>	<i>Boys.</i>		<i>Girls.</i>	
	<i>Number.</i>	<i>Mean Mental Age.</i>	<i>Number.</i>	<i>Mean Mental Age.</i>
2 to 2; 11	20	2· 7	15	2· 6
3 to 3; 11	20	3·58	20	3·46
4 to 4; 11	25	4·49	16	4·47
5 to 5; 11	39	5·57	24	5·04
6 to 6; 11	28	6·38	37	6·22
7 to 7; 11	44	7·39	50	6·72
8 to 8; 11	35	9·16	46	8·16
9 to 9; 11	39	9·79	41	8·94
10 to 10; 11	61	11· 3	42	9·67
11 to 11; 11	55	11·55	38	10·65
12 to 12; 11	54	12·67	31	12·58
13 to 13; 11	50	14· 0	32	14·06
14 to 14; 11	38	14·53	18	13·73
15 to 15; 11	44	15·86	13	14·83
16 and above	86	15·75	13	14·83

It will be seen that the boys and girls start together, but as they grow older the differences between them become wider and wider, the girls throughout progressing at a lower level until about the eleventh year, when their curve again shows a steeper ascent and overtakes that of boys at about the thirteenth year. They again fall back slowly at higher ages.

Baldwin and Stecher¹ tested 143 American children repeatedly by the Standard Revision of the Binet Test. Comparing the Indian results with one of Baldwin and Stecher's graphs we see that there is a somewhat

¹ BALDWIN, B. T., and STECHER, L. I.: *Mental Growth Curve of Normal and Superior Children, Iowa Studies in Child Welfare*, Vol. II, No. 1, 1922, pp. 10-11. Compare also F. D. BROOKS: *The Psychology of Adolescence* (Harrap), Chap. V.

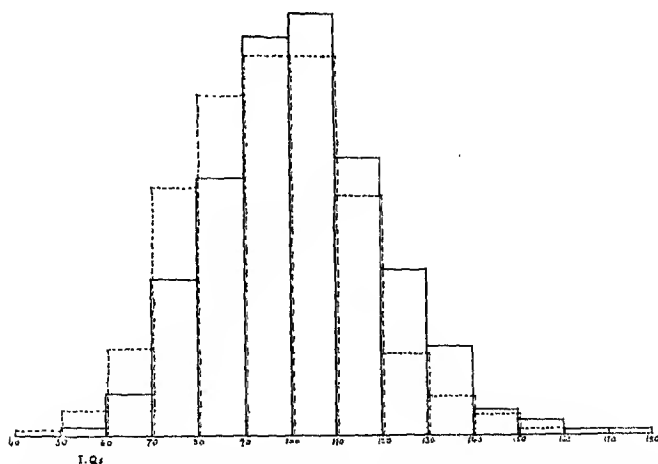
close resemblance between the two graphs of growth, if some allowance be made for the different environment of Indian girls. In both there is a tendency for the graph of girls' growth to be slightly below that of boys until about the eleventh year. From the eleventh to about the fourteenth the girls' graph rises more steeply and is higher than that of boys in Baldwin and Stecher's diagram and reaches that of boys in the Indian study. Finally, the girls' graph falls back once more. The steeper rise of the girls' graph between eleven and fourteen may be explained by their earlier adolescence as compared with boys.

Professor Burt in his London Revision found, on the other hand, that the mental ages of girls at almost all ages were slightly higher than those of boys. He says: "Sheltered, supervised, detained at home, girls, like children of the better classes, incline to sedentary lives and engage in literary pursuits; and, like those children, they consequently excel in linguistic work and conversational activities."

There is no doubt that environment, social or otherwise, plays a great part in the performances of children in all mental tests. If as Burt says the English girls being sheltered, supervised and detained at home incline more towards literary pursuits, the Indian girls, on the other hand, suffer from a comparative neglect of their literary education.

III.—COMPARISON OF HISTOGRAMS OF BOYS' AND GIRLS' I.Q.S.

In the present Indian experiment the number of boys tested was 638 and that of girls 436. The distributions of their I.Q.s raised to 1,000 in each case for easy comparison are given by the accompanying histogram:



The unbroken line Histogram belongs to boys and the broken line one to girls.

Looking carefully at the histogram we see that the distribution of girls is shifted slightly towards the left, while the reverse is the case with boys. This may be due to some extent to the superior educational environment of boys as stated above.

The mean I.Q. of all the boys of all ages put together was 102.29, and the corresponding I.Q. of girls was 96.12. The difference of 6.17 is statistically significant, since the standard deviation of the difference equals 1.13 and $\frac{\text{Difference}}{\sigma \text{ Diff}} \approx 5.46$. The standard deviations of the two distributions were 18.46 in the case of boys and 18.0 in the case of girls. The scatter of the girls' distribution is thus almost the same as that of the boys. The suggestion made by some that girls cluster together more at the centre scarcely holds good in this case.

IV.—COMPARISON OF BOYS' AND GIRLS' PERFORMANCES IN INDIVIDUAL TESTS AND IN PREDOMINANT ABILITIES.

The percentages of boys and girls separately that passed in the various tests of the scale were then calculated. Here it was noticed that in the lowest group of tests, namely, those of year III, the boys had three items to their credit and the girls four, and the differences in the performances of the two sexes were insignificant. As we ascend the scale into the higher age groups the number of items in which girls show superiority quickly diminishes until from the eighth year group onwards the girls disappear altogether, the boys showing a superiority in all the tests. The differences in their performances also become more and more significant. These differences again become smaller from the twelfth year group and in the fourteenth year group the girls actually show superiority in three of the tests. Further on they begin to fall back again. Thus the effect of earlier adolescence of girls commencing about the eleventh or twelfth year is seen in their performances in individual tests also.

All the tests in the scale were then grouped in a naive manner into eight predominant abilities and the number of items captured by the two sexes in these abilities was noted. The following table gives the results :

<i>Kind of Ability.</i>	<i>No. of tests in which boys were superior.</i>	<i>No. of tests in which girls were superior.</i>	<i>No. of Significant Differences.</i>
(1) Language or verbal ability	15	1	3 (in favour of boys)
(2) Arithmetical ability.....	6	—	1 (in favour of boys)
(3) "Performance".....	3	3	1 (in favour of boys)
(4) Immediate memory	11	3	2 (in favour of boys)
(5) Æsthetic and drawing ability	6	4	Nil.
(6) Comprehension.....	6	—	Nil.
(7) Imagery and association ..	8	1	Nil.
(8) General knowledge	15	1	1 (in favour of boys)

In 'performances,' that is, in manipulating ability, the boys and the girls have an equal number of items to their credit. The girls are also nearly as good as the boys in æsthetic and drawing ability. They show well also in the immediate memory tests of lower ages. In all other abilities the boys show decided superiority. In arithmetical ability and in comprehension the girls show marked inferiority. It may be noticed that in language or verbal ability, in which the feminine sex is supposed to excel, the boys have shown themselves as decidedly superior, and in performance, where one would have supposed that boys would do much better, the two sexes have shown themselves to be equal. Thus the stories of a "linguistic sex" or a mechanically-minded sex would seem to be myths. The superiority shown by the sexes in such abilities would seem to depend on the opportunities they get to develop these abilities in their daily routine.

V.—SUMMARY AND CONCLUSIONS.

Absence of agreement in the results of mental measurements of different workers seems to be partly due to the indefiniteness of the unit of mental measurement and the nature of the measurement itself. It is also partly due to the fact that mental abilities are more susceptible to change by environment and effort than physical abilities. Again, it must be remembered that man puts forth far greater labour in order to train the mind than he does to train the body. Thus it is quite possible that this immense labour in mental work may wipe out the small native differences in endowment. We are not agreed also what exactly intelligence is or what is the best method of measuring it. Does the ability of the mind depend upon the brain and the nervous system? If we postulate that it does the next question is does intelligence or mental ability depend entirely on the quality of the nervous system or has it anything to do with the quantity also? If it depends only on the quality then there ought not to be any difference between the sexes as the quality of the germ-plasm would be the same in either case. If the quantity also has to be taken into account we may expect some difference in the mental abilities of the two sexes. At any rate, our methods of measuring mental abilities depend on measuring the output of work done by the mind in a definite amount of time. In other words, to borrow an idea from mechanical science we measure the *power* of the mind. If this be so may it be that if boys and girls were reared up in exactly identical conditions the mental growth curves of boys and girls would show the same characteristics as their physical growth curves? The slightly higher norms for girls than expected obtained by Burt could then be explained by the fact that girls in England incline towards literary pursuits being

"sheltered, supervised and detained at home."¹ On the other hand, the slightly lower norms of Indian girls would be due to their comparatively inferior literary environment.

Restricting ourselves to experimental studies it is evident that environment and effort play a great part in determining mental abilities. Studies of sex differences and social differences make this quite clear. Taking all studies on sex differences together it would seem that the mental growth of the two sexes is very nearly on the same level; the slightly superior ability shown by either sex may be due to environmental effects. At any rate, environment and effort play such a great part in mental abilities that even if there be any small native differences they are swamped under the influence of these. The slightly greater positive skewness of the boys' histogram in this Indian study makes it probable that there has been some effect of different environment. There are indications also that the curves of physical growth and mental growth resemble each other in particular in respect of bringing adolescence earlier in the case of girls, and fixing the limit of growth at adulthood in the case of both sexes.

¹ BURT, CYRIL: *Mental and Scholastic Tests* (P. S. King and Son, 1927).

THE EFFICIENCY OF DIFFERENT METHODS OF MARKING ENGLISH COMPOSITION.

By B. M. D. CAST.

(From the Department of Psychology, University College, London).

PART I.

I.—*Problem.*

II.—*Experimental and statistical procedure.*

III.—*Variations in marks.*

VI.—*Analysis of variance.*

I.—PROBLEM.

"THE modern examination," it has been said, "is dominated by the essay. It is based on the essay; it is built of the essay; it stands or falls by the measurability of the essay. And by the essay is to be understood anything that may be called English composition—any attempt to express ideas in discursive prose."¹ This traditional faith in the essay type of question has been partly justified, and partly shaken, by systematic investigations upon scholastic tests. In the earlier inquiries the correlations between tests showed plainly that of all forms of school work English composition had the closest relation with intelligence and with general educational ability.² At the same time, the correlations between persons—i.e., between the different examiners' markings of the same test-performances—indicated that the assessment of essays was one of the most unreliable modes of mental measurement. "No other form of examination," says Burt, "leads to such inconsistent marking." "In all subjects into which the factor of composition enters," writes Professor Valentine, "and in the marking of essays more especially, extraordinary variations occur between the marks of different examiners."³ The reports of the investigators of the School Certificate Examinations and of the International Examinations Inquiry Council come to the same conclusion: "the English essay is . . . notoriously difficult to mark"; "the common methods of marking pupils' compositions are from a scientific standpoint almost worthless."⁴

¹ BALLARD: *The New Examiner*, 1923, p. 52.

² BURT: *The Distribution and Relations of Educational Abilities*, 1917, p. 55. Cf. *Mental and Scholastic Tests*, 1921, p. 330.

³ BURT, *loc. cit.*; VALENTINE: *The Reliability of Examinations*, 1932, p. 26; cf. p. 167.

⁴ *The School Certificate Examination*, H.M. Stationery Office, 1932, p. 37. HARTOG and others: *Marks of Examiners*, 1936, pp. 117 *et seq.*

The moral drawn by most educational psychologists can be summed up in the phrase which heads Ballard's chapter on the subject—"the rejection of the essay": "in any other science an instrument so imperfect would have been cast on the dust-heap." But there is an alternative; instead of discarding it, we may seek to improve it.

The fact is that neither teachers nor examiners have ever systematically inquired what are the most accurate and effective methods of marking. The need for a comparative study of such methods was strikingly shown by an early inquiry carried out in London in connection with the examination for junior county scholarships. It was found that:

"Quite apart from the extraordinary differences in the average, range, mean variation, and frequency curves of the marks themselves, the methods and criteria employed vary widely: at least fourteen different kinds of procedure may be enumerated . . . Most of these may be reduced to variants either of the impressionistic principle or of the analytic principle, or an unequal combination of the two. . . . There is the widest variation in the weight allotted to different aspects, particularly to the more superficial and mechanical processes (where inaccuracy at once attracts the notice of the impressionistic examiner) as contrasted with the higher and more organized processes of style and logic (nearly always undermarked). Where the method is professedly analytic, the underlying scheme is never systematic and comprehensive, and is seldom based on any logical or psychological plan. Nor indeed have we as yet any verified knowledge as to which of the proposed devices is really the most effective, or (since here as elsewhere there are doubtless wide individual differences) which method is best suited for this type of examiner and which for that, or indeed what types of examiner there are."¹

II.—EXPERIMENTAL AND STATISTICAL PROCEDURE.

The inquiry described in the following paper was designed primarily to seek some solution to these problems—to investigate the reliability and validity of the chief principles of marking actually employed, and to see whether there are any broad or typical individual differences in the efficiency with which each is used by different examiners.

Forty compositions were obtained from forty central school girls aged between fourteen and a half and fifteen and a half. The topic was one suggested by Sir Philip Hartog²: "'Depend in life only on your own efforts': discuss (a) how far such a course is possible in a civilized

¹ BURT: *Report on Junior County Scholarship Examinations*, 1917, p. 24.

² Cf. HARTOG: *The Writing of English*, 1908, p. 142. The general problem and methods were suggested by Professor Burt, whose help at all stages of the inquiry I gratefully acknowledge. I am also much indebted to Sir Philip Hartog and to Dr. Ballard for allowing me to discuss with them the procedure that might best be followed in obtaining typical compositions from pupils in school and typical specimens of marking from examiners. I am most deeply indebted to the various masters and mistresses who so generously gave their time and attention to the task, and co-operated so willingly in the whole experiment.

country, and (b) how far it is good." The subject was announced two days beforehand. One hour was allowed for writing. Typed copies of the children's compositions, preserving the original spelling, punctuation, and paragraphing, were then submitted to twelve judges—five head teachers and seven assistant teachers in London schools, all experienced examiners and all keenly interested in the experiment. Subsequently, in order to obtain a mark for handwriting, the original essays were themselves circulated.

Each examiner was asked to mark these essays according to the four following methods. The more obvious precautions (cyclic order, long intervals between the markings, etc.) were taken to prevent memory of earlier marks affecting the later. Both the statements of the examiners themselves and the "self-consistency coefficients," i.e., the correlations between the marks awarded on successive occasions by the same examiner ("reliability coefficients," as they would usually be termed¹), afford some evidence that these precautions were reasonably successful. The work was planned and started in 1935; but the long intervals between the markings and re-markings have delayed its completion.

(1) *Individual Method*. The examiner was first asked to mark the compositions according to the method that he himself would ordinarily adopt, and to indicate the principles followed, the scale of marking adopted, and the chief criteria he had observed. Five out of the twelve relied mainly on some method of general impression; three on some form of detailed analysis; four adopted a combination of both (cf. Table VII). For comparison all marks were reduced to a percentage scale.

(2) *Achievement of Aim*. After an interval of about eight weeks, each examiner was asked to mark the same compositions again, this time following the principle suggested by Sir Philip Hartog. He argues that in writing a genuine composition the author is addressing some particular audience, and has some particular object in view—to instruct, persuade, or entertain the reader, or possibly to work out in writing the answer to a problem that he has set himself. The proper test should, therefore, be: "how far has the writer achieved the purpose which he has implicitly put before himself?"² With this method the examiners were required to mark on a percentage scale.

¹ In such work as the present it is necessary to distinguish, as Burt has pointed out, between two kinds of so-called reliability coefficients: the one kind he calls "internal, intra-individual (or self-) consistency coefficients," the other "external (or inter-individual) consistency coefficients." In form both may be "correlations between persons, but with the former there need only be *one* person." Here, for example, we are concerned with correlations between different markings carried out by one and the same examiner. I shall refer to the interesting results obtainable by factorizing these "self-consistency coefficients" later. For the moment it is sufficient to note that, even on the fourth ranking, the correlations were not high enough or different enough to suggest that memory or practice had influenced the later markings appreciably or differentially; the point is important because, as it turned out, the method used last of all furnished the best results.

² See HARTOG: *The Writing of English*, 1908, pp. 60-2. Cf. *Id.*, "English Composition," *Essays on Examinations*, 1936, pp. 135-8.

(3) *General Impression*. After another interval of about two months, the examiners were asked to mark the compositions once more by general impression. Hartog and Rhodes, from their studies of marks allotted to English essays in examinations like those for special places, conclude that "the discrepancies between marks awarded by impression and marks awarded by totalling for details . . . are entirely due to the greater difference in the standards of marking of different examiners when they mark by impression" (i.e., to their different averages).¹ Accordingly, to eliminate these influences, it was decided that the same average and the same type of distribution should be adhered to by all. The instructions advised that the forty essays should be first grouped into five broad classes, then arranged so far as possible in an order of merit, and finally allotted marks on a scale from 0 to 100 conforming with a normal distribution: the marks recommended were based on Burt's "standard scale."

(4) *Analytic Method*. The procedure which involved the most intensive study of the scripts was kept until last. With this method the examiners were requested to allot marks separately for each of the main aspects or elements of a good English composition, and to base their final marks on the total. The schedule of items was that worked out by Burt, and summarized in his *Report*.² Briefly it includes not only (1) the more mechanical aspects of composition—handwriting, spelling, punctuation, grammar, syntax, but also (2) the more strictly literary aspects of composition—range, correctness, and appropriateness of information, of vocabulary, and of rhetorical devices, and, above all, (3) the logical aspects of composition—that is, the general organization of ideas, as revealed by the unity, the complexity, the relevance, and the sequence of sentences and of paragraphs, and by the intellectual structure of the essay as a whole. Detailed notes appended to the schedule indicated what maxima were to be awarded under each heading and the principles on which the specialized marking is to be based.⁴

To study the relative efficiency of these four methods, three main procedures have been followed: first, the old and simple method of comparing averages and random deviations; secondly, the comparatively new method of "analysing variance"—a method devised originally for agricultural research by Fisher and adopted for such problems as the present by Burt; thirdly, the so-called "P-technique"—an adaptation of the usual methods of factor-analysis for dealing with correlations between persons instead of between tests or traits. Since this is the first research in which all three procedures have been employed on the same set of data, it may be regarded as a study of methods quite as much as of results. Its specific object was to test, by an *ad hoc* experiment, the conclusions drawn by Burt by analysing actual examination results.

¹ *Marks of Examiners*, p. 123.

² 1917 *Report*, p. 49, *Backward Child*, p. 26.

³ *Mental and Scholastic Tests*, p. 331.

⁴ For fuller details regarding the methods and results I may refer to my thesis on *An Investigation into the Methods of Marking English Composition* (University of London Library).

III.—VARIATIONS IN MARKS.

The time-honoured method of analysing such data is twofold: "we look (i) for constant differences between examiners' standards as indicated by differences between their averages, and (ii) for random variations as indicated by the discrepancies remaining after allowance has been made for differences of standard."¹ Since the mark chosen as indicating the average standard is perfectly arbitrary and can be readily allowed for (or so at least it is commonly supposed) by "equating averages," the best measure of inefficiency is generally taken to be given by the amount of "random variation," i.e., by the general amount of "deviation or divergence from the ideal mark allotted by general consensus to each candidate." This is variously assessed by calculating the extreme range, the average range, the mean variation, or (in most recent inquiries) the standard deviation. Thus, it is claimed, "we may classify our examiners according to these standard deviations."²

TABLE I.
DIFFERENCES IN GENERAL STANDARD.

Method.	Lowest Average.	Examiner.	Highest Average.	Examiner.	Difference.
Individual	46.1	E	67.7	J	21.6
Achievement	42.2	A	75.6	F	33.4
General Impression	47.6	J	55.1	L	8.5
Analytic	45.7	D	61.8	G	6.1

(i) The differences of standard are shown in Table I. Where examiners are left to their own devices they vary in the averages they award to the group by about 20 per cent of the maximum: this agrees very closely

¹ E. C. RHODES: *Marks of Examiners*, p. 185. The application of these two methods to such problems as the present is usually attributed to American psychologists, but in this country it goes back to Edgeworth's early papers (*The Statistics of Examinations*, *J. Roy. Stat. Soc.*, LI, 1888, pp. 599-635, and *The Element of Chance in Examinations*, *ibid.*, LIII, 1890, pp. 460-75, 644-63).

² RHODES, *loc. cit.*, p. 193. Actually in their study of the marking of the English essay, Rhodes and Hartog more frequently rely upon "range of marks." In a memorandum appended to their volume Burt had suggested the use of factor analysis and of Fisher's methods. Rhodes, however, in a series of unpublished memoranda, argues strongly against the application of correlational methods; but it is not quite clear why he rejects the proposal to apply the analysis of variance, particularly as his illustrative table (117, p. 187) seems backed on this principle and the small size of his groups would seem to make Fisher's formulæ especially appropriate. He makes, however, the interesting suggestion that the discrepancies should be calculated from a weighted average, not the ordinary average—a proposal which corresponds to the substitution of factor-analysis for the ordinary analysis of variance.

with the findings of previous investigators. Where Hartog's criterion is adopted ("how far does the writer achieve his purpose?") they vary still more: here, adopting the borderlines for failure and distinction adopted in many London examinations, we may say that A would regard nearly half the children as "failures," while F would regard more than half as meriting "distinctions"! The reduction of differences with the last two methods is due to the instructions, which required all examiners to adhere to the same average—directly in the case of the method of General Impression, and indirectly in the case of the Analytic method. In view of the objections so often raised against the "artificial restriction" of the general average by explicit instruction, it is interesting to find that the Analytic method *indirectly* leads to an even closer equalization of standards.

TABLE II

MARKS AWARDED TO THE SAME CANDIDATES BY DIFFERENT EXAMINERS.

<i>Method.</i>	<i>Extreme Range.</i>	<i>Average Range.</i>	<i>Standard Deviation.</i>
Individual	74	47.1	12.5
Achievement	75	53.5	11.3
General Impression ..	75	44.6	12.4
Analytic	56	34.9	8.8

(ii) These differences in standard, however, do not affect the order of the candidates. A more important sign of inefficient marking, therefore, is the wide range or variability¹ shown by the marks allotted to one and the same script, even after the differences in general standard have been eliminated. The chief figures are summarized in Table II. Judged by this criterion, the Analytic method would seem by far the best out of the four. With the Individual method the extreme range is 74 marks out of 100 (this occurred twice—Examiner C giving 89 marks to Script 16 and E only 15, Examiner J giving 80 to Script 30 and C only 6). With the methods of Achievement and General Impression the range is practically the same (75 marks—a divergence only found once in each case). With

¹ The first use of the variability as a measure of the unreliability of a set of psychological judgments is usually ascribed to J. McK. Cattell. He proposed to rule out differences of standard by using ranks. The most elaborate study with Cattell's method is F. L. Wells's early investigation on *The Variability of Judgments*—a case where a correlational method would now almost certainly be employed (*Essays in Honour of William James*, 1908, pp. 509–550). The formulæ and references are given in Woodworth's recent *Experimental Psychology*, 1938, pp. 372 *et seq.*

the Analytic method the extreme range is greatly reduced (namely, to 56 marks). If we take the standard deviation for our measure, we may say that the other methods appear to increase the amount of disagreement by nearly 50 per cent over that produced by the Analytic method.

These results are not so conclusive as might at first sight be supposed. Others have noted that, with impressionistic methods, the range of marks awarded by different examiners to one and the same script tends to be unusually wide; and it has been inferred that such wide discrepancies alone suffice to demonstrate the unreliability of such methods. That is by no means a necessary conclusion; for it is conceivable that the averaging involved in the analytic methods may tend to reduce the range of the final marks unduly, and so obscure differences between candidates that the impressionistic method sometimes preserves.

Let us, therefore, compare the ranges, not for different examiners with the same candidates, but for the same examiner with the different candidates. The figures are given in Table III. With the Analytic method, the extreme range, the average range, and the standard deviations for nearly all the examiners are much lower than with any of the other methods. The method of General Impression spreads the candidates most widely. But with this method the curve of distribution is somewhat abnormal. The less experienced examiners try to spread the mediocre candidates too much and do not spread out the extreme cases sufficiently.

TABLE III.

MARKS AWARDED BY THE SAME EXAMINER TO DIFFERENT CANDIDATES.

<i>Method.</i>	<i>Extreme Range.</i>	<i>Average Range.</i>	<i>Standard Deviation.</i>
Individual	87	60.1	15.76
Achievement	87	60.2	16.30
General Impression ..	80	71.8	17.83
Analytic	76	52.9	12.79

In the results of the Analytic method we are evidently confronted with an important but neglected tendency which may be noted in other types of examination, viz.: the "reduction of range by averaging." "If an entire script, containing say half a dozen answers, is marked by general impression, the standard deviation is bound to be larger than if the examiner marks the six component questions separately and then averages his marks. To expect anything else is to assume a perfect correlation between the marks for individual questions. A candidate who gets the highest mark in one particular question will almost certainly get lower marks than the highest in the remaining questions; similarly with the candidate who gets the lowest mark. Thus, on averaging their marks for all the questions, the range inevitably shrinks. This "regression" is the inevitable consequence of all forms of

summation of incompletely correlated figures."¹ What holds of averaging marks for separate answers holds also of averaging marks for separate points in one and the same answer or essay.

A first-hand study of the detailed marks, side by side with the scripts themselves, reveals further differences. The impressionistic methods, for example, appear specially liable to seize on a few salient or superficial points—errors of spelling, grammar, or fact, perhaps—and weight those out of all proportion to the rest: on the other hand, the analytic methods, by dealing with numerous isolated and possibly inessential points, may overlook certain general qualities that characterize the essays as a whole. These tendencies can only be corrected by some kind of differential weighting. It seems clear, therefore, that a factor-analysis is required to show what aspects should be most heavily weighted and which deserve such small weights that they might be safely dropped.

Meanwhile, the simple and time-honoured procedure of contrasting averages and variabilities does not by itself prove that the differences are genuinely significant or determine whether they are greater or less than in other researches. Are we to calculate the probable error for each difference? With twelve examiners and four modes of marking there would be 1,128 differences to test. Clearly we need some more condensed and intensive method of analysis; and this is provided by an analysis of the "variance,"²

IV.—ANALYSIS OF VARIANCE.

Hitherto but little use has been made in psychology of this procedure. Burt has adopted it both for the study of mental differences between individuals³, and for the study of junior county scholarship, teacher's diploma and certificate examinations.⁴ There seems, however, to be considerable doubt how far it is practicable to "analyse variance" in psychological work. Stephenson, for example, in criticizing the proposal, insists that it would be quite illegitimate with psychological data:

¹ BURT: *Memorandum on the Technique of Marking*, p. 23. The practical importance of this consideration will be clearer if I quote the succeeding paragraphs: "What holds good when question-marks are summed for each single script by one examiner will also hold good when marks for different papers are summed by a Board. And it will still hold good even when the maxima or the averages for the different papers have been equalized. A common fallacy is the following. A mark is prescribed for failure or for distinction—say one which will cut off the top and bottom 2 per cent. This mark refers to the averages of all papers. Each examiner, however, adopts the same borderlines in marking individual scripts. The natural result is that on summation hardly anyone fails and hardly anyone scores a distinction. For plainly, if only two candidates out of a hundred score over 80 per cent on each of six papers, these will not be the same candidates. Hence 2 per cent cannot score over 80 per cent on the totals."

² The variance is the square of the standard deviation. Unlike the standard deviation, variance is additive, and can therefore be separated into component parts.

³ Cf. "The Distribution of Temperamental Types" (*Factors in the Mind*, in the press).

⁴ L.C.C. *Report*, loc. cit. sup. His arguments are briefly summarized in *Brit. Jour. Psychol.*, XXVIII, i, p. 61, and *Proc. Roy. Soc., Ser B*, CXV, p. 419, "Recent Developments of Statistical Method" (reprinted more fully in *Occupational Psychology*, XII, iii, p. 170).

"variance in psychology," he maintains, "depends for its size merely on the whims of the psychologist." Accordingly, one of the objects of the present inquiry was to study how far the proposed procedure would be applicable and fruitful.

Burt's earlier analyses of actual examination results were intended rather to illustrate the nature and feasibility of the procedure than to reach immediate and concrete conclusions.¹ Marks awarded in real examinations do not, as a rule, conform very precisely to the ideal conditions of a statistical inquiry (e.g., for a satisfactory "analysis of variance" *all* the examiners should mark *all* the scripts, and that in circumstances which eliminate constant or progressive influences so far as possible; this is seldom possible in practice). For this reason, therefore, among others I decided to work rather with artificial data, that is, with essays obtained and marked expressly for the purposes of scientific study.

We have seen in the previous section that the total variation between marks awarded to different candidates by different examiners may conceivably be due to one or both of two main sources: variation may arise either from real differences in merit between the candidates or from subjective differences in the examiners' power to estimate that merit. The latter source of variation in turn is twofold: it depends partly on the differences in the standards adopted by the several examiners and partly on the differences in their ability to eliminate random errors.² Thus the "total variance" (V_t say) can be split up into three component parts: (i) a variance due to differences in merit between candidates, i.e., to variation in the average mark allotted to each candidate (V_c); (ii) a variance due to difference in the standard adopted by the several examiners, i.e., to variation in the average mark that each examiner tends to award (V_e); and (iii) a residual variance due to random errors (V_r).³ The object of "analysis of variance" is simply to estimate the relative size of these three components, and, taking the third (the residual variance) as a basis of comparison, to inquire whether the first and the second are statistically significant as compared with the size of this error.

¹ The memoranda based on actual data obtained from certain examinations carried out by the London County Council or the University are confidential; and I gratefully acknowledge the permission to refer to them in general terms: (it would be improper for me to quote explicit results that seem to confirm my own, since such quotations might be construed as reflecting on the efficiency of official examiners).

² See *Marks of Examiners*, Memorandum I, pp. 260 *et seq.*

³ For those who are more familiar with a correlational factor-analysis, it may be pointed out that, if we regard each examiner as a living "test," the differences between candidates' averages (V_c) would in factor-analysis be attributed to a "general factor for tests" (i.e., to a factor based on the [weighted] average of all the "tests" or examiners) whereas V_e would be attributed to a "general factor for persons" (i.e., to a factor based on the [weighted] average of all the "persons" or candidates). The residual variance is a doubly centred bipolar matrix such as that described by Burt in extracting identical secondary factors for tests and persons.

266 *Efficiency of Methods of Marking Composition*

It has been said that "the two initial requirements of a good examination procedure are (1) that it should differentiate as significantly as possible between the candidates, and (2) that it should not differentiate significantly between the examiners who use it"¹—i.e., that it should reveal the candidates as a significantly heterogeneous group and the examiners as a homogeneous (or at least an insignificantly heterogeneous) group. Let us then apply these two criteria to the present data. To express them in terms of a single figure, I shall take the ratio of (i) the variance for the candidates (V_c) and (ii) the variance for the examiners (V_e) to the residual variance (V_r).

TABLE IV.
ANALYSIS OF VARIANCE.

Method.	Source of Variation.	Degrees of Freedom.	Sum of Squares.	Mean Square.	F-Ratio.
Individual ..	Between Examiners	11	19859.46	1805.41	11.62
	Between Candidates	39	61132.49	1567.50	10.09
	Error	11 × 39	66623.88	155.30	
	Total	479	147615.83	308.18	
Achievement	Between Examiners	11	56813.72	5164.88	31.95
	Between Candidates	39	69329.53	161.61	9.81
	Error	11 × 39	69329.53	161.61	
	Total	479	187984.90	392.45	
General Impression	Between Examiners	11	2640.57	240.05	1.39
	Between Candidates	39	76818.30	1969.70	11.42
	Error	11 × 39	73938.60	172.35	
	Total	479	153397.47	320.24	
Analytic	Between Examiners	11	12221.48	1111.04	12.26
	Between Candidates	39	41605.09	1066.79	11.77
	Error	11 × 39	38867.36	90.60	
	Total	479	92693.93	193.52	

¹ BURR, *loc. cit. sup.*

The results of the analysis¹ are set out in Table IV. The last column but one gives the "variances" or mean squares, calculated by using "degrees of freedom" instead of the number of items summed. The last column of all gives the ratios of the two main variances to the error variance. The published tables² show that with 11 and 39 degrees of freedom for the larger variance, and 11×39 for the smaller, the 5 per cent point of significance should be 1.81 and 1.42 respectively, and the 1 per cent point 2.29 and 1.64. Only one of the ratios, therefore, fails to be significant—that for the method of General Impression: but here, it will be remembered, practically the same standard was imposed on all examiners by the instructions. With this artificial exception, therefore, we may conclude that both the examiners (judged by their choice of mean) and the examinees (judged by the averages allotted to each) form heterogeneous groups in the sense above defined.

TABLE V.

SUCCESS OF THE SEVERAL METHODS.

- in (i) Differentiating between candidates (V_c/V_r).
 (ii) Reducing differences between examiners (V_e/V_r).
 (iii) Eliminating random error (V_r/V_t).

	(i) V_c/V_r	(ii) V_e/V_r	(iii) V_r/V_t
Analytical Method	11.77	12.26	0.47
Method of General Impression	11.42	1.39	0.54
Individual Methods	10.09	11.62	0.50
Achievement Method	9.81	31.95	0.41

¹ To describe in detail the actual process of calculation is unnecessary. The general principles are given in FISHER: *Statistical Methods for Research Workers*, 1934, pp. 198–235; and SNEDECOR: *Statistical Methods*, 1937, pp. 179–248. Owing to the peculiar nature of most measurements in psychology, and particularly of marks allotted in examinations, some re-adaptation is necessary, and certain special precautions have to be observed before the method can be used for such problems as the present. Here I have followed the procedure used by Burt in analysing the variance of marks awarded in the examinations just referred to, described in his *ronoe'd Notes* (to be obtained from the Department).

² SNEDECOR, *loc. cit.*, p. 187.

268 *Efficiency of Methods of Marking Composition*

On the basis of these figures we may attempt to arrange the four methods of marking in an order of merit. Three principles of comparison may be used.

(1) First of all, the greater the variance between the means for the candidates, as compared with the random "error," the greater and the more trustworthy is the differentiation between the several candidates. Hence the methods may first be ranked according to the extent to which they secure this differentiation (Table V, col. i). The most successful has been placed first. Judged by success in differentiating between the candidates, the Analytic method seems best. The method of General Impression is almost as good. But when allowed to follow their own favourite method, examiners were somewhat less successful. The Achievement method appears the poorest of all.

(2) When we change our criterion and consider the degree to which the methods introduce differences between the different examiners, then the method of General Impression might at first sight seem the most satisfactory ; but, as we have seen, the negligible difference is here due mainly to the fact that the initial instructions virtually imposed the same mean on all. The wide variation produced by the Achievement method is presumably due to the disagreement about what can constitute a 50 per cent degree of success. With a set of arithmetical problems most examiners would almost inevitably agree over the aim which the examinee was setting himself and over the success with which his aim was fulfilled ; but with essays, as the examiners' subsequent comments reveal, there is a wide divergence both over the aim which they impute to the various writers and over the subjective reactions which they would expect from the supposed audience or reader. I feel, therefore, that, although Hartog's criterion forms an admirable principle for teaching, it is too precarious to form the *sole* principle for marking.

(3) Thirdly, the extent to which the different methods have enabled the examiners as a group to reduce random fluctuation in their marking may be assessed by comparing what is termed the "variance of the error" with the "variance of the total" (i.e., the variance calculated about the mean of the whole sample). When we judge by this last criterion, we find that, of all four methods, the method of General Impression shows the greatest amount of random fluctuation. For this there seem to be three reasons. (i) The fact that each examiner was using his own criteria must occasionally introduce peculiarities into his judgments, which, compared with those of the remainder, may appear to be erratic. (ii) In a few cases the fact that the instructions imposed the same wide variation on every examiner, whether or not he was prepared to make it spontaneously, must have exaggerated the effects of minor factors. (iii) But undoubtedly the main cause lies in the degree to which all subjective impressions are influenced by chance elements and unconscious caprice.

The Analytic method, on the other hand, tends to ensure that even small variations shall have objective causes. What, however, is most surprising, and by no means easy to explain, is that the Achievement method shows the smallest ratio of all. Possibly the small size of its ratio is due, not so much to the small figure for the random error, as to the large figure for the total

variance. These, however, are among the questions that cannot be answered by an analysis of variance alone.¹

¹ Before leaving the subject of analysis of variance, I should like to emphasize one valuable application which I have no space to illustrate at length. In the foregoing research all the children were from the same school ; but, though some of the examiners were from the same school and knew the children already, others were from different schools. This is by no means an uncommon situation. It is, for example, one of the issues discussed at length in the *Memorandum* above referred to : when pupils or students from a particular school or college are marked by their own teacher and obtain from him higher marks than he awards to those from other institutions, it is often asked : has he not unconsciously favoured the pupils whom he knows or perhaps the type of answer he has inculcated ? The teacher usually replies that the high marks are due to the fact that his school or college draws a better type of pupil. A twofold analysis of variance will nearly always show whether the main source of such differences is due to the examiner marking or to the institution from which the marks are obtained. I understand that this principle has already been tried in certain examinations for intending teachers : it would seem equally applicable to examinations of school pupils. This, however, is only one of the many pressing practical problems which might be solved by an analysis of variance.

Part II of this article will appear in the next issue. It contains ;

V.—*Factor analysis.*

VI.—*Types among examiners.*

VII.—*Types among examinees.*

VIII.—*Conclusions and summary.*

PRIMARY MENTAL ABILITIES.

(PSYCHOMETRIC MONOGRAPHS No. 1.)

By L. L. THURSTONE. (Chicago: University of Chicago Press, pp. x+121. 9s.)

THIS publication is the opening number of a series which the Psychometric Society proposes to issue. It reports the first large experimental inquiry, carried out by the methods of factor analysis described by Thurstone in *The Vectors of the Mind*¹. The work was made possible by financial grants from the Social Science Research Committee of the University of Chicago, the American Council of Education, and the Carnegie Corporation of New York. The results are eminently worthy of the assistance so generously accorded. Thurstone's previous theoretical account, lucid and comprehensive as it is, is intelligible only to those who have a knowledge of matrix algebra. Hence his methods have become known to British educationists chiefly from the monograph published by W. P. Alexander². This enquiry has provoked a good deal of criticism, particularly from Professor Spearman's school; and differs, as a matter of fact, from Thurstone's later expositions. Hence it is of the greatest value to have a full and simple illustration of his methods, based on a concrete inquiry, from Professor Thurstone himself.

Fifty-six tests, selected according to a provisional classification of cognitive factors, were applied to 240 volunteers. The correlations between the tests were then estimated by means of the charts for tetrachoric correlation previously published by Thurstone and his colleagues.³ The huge table of correlations has been factorized by the so-called centroid method; and twelve factors extracted. All except the first are bipolar, i.e., have negative as well as positive saturations. It is then assumed that "primary factors act positively unless they are absent from a performance." Hence the co-ordinate axes, representing the factors, are rotated, two at a time, until the negative saturations are virtually obliterated, and the number of zero saturations maximized. For this purpose thirty diagrams have been plotted, and fresh axes fitted by eye.

It is stated that "the graphical method of rotating in one plane at a time is probably the best single method; but the graphical method is not ideal." Although the principles involved have been briefly explained in *The Vectors of the Mind* and elsewhere, hitherto, as Professor Thurstone points out, "there has not been published any adequate description of the method as applied to an actual problem." This,

therefore, is in some ways the most interesting section of the report. The final upshot is thirteen fresh factors, of which nine can readily be given a psychological meaning. In the main, though not in every detail, the interpretations correspond with the categories which the tests were originally selected to represent.

When the editor of this *Journal* first suggested a review of Professor Thurstone's new report, it appeared that the large collection of data contained in its tables would offer an admirable opportunity for testing recent statements about the mode of factor analysis, statements for the most part reached *a priori* and never yet verified by any concrete comparison. How, for example, do Thurstone's methods and results compare with earlier methods and results put forward by workers in this country?

In his 1935 Memorandum (³, page 306) Burt has pointed out that there are in theory two general ways of factorizing a table of correlations between tests: (a) with the first method—a 'submatrix' or 'group factor' method—we may look for relatively specific factors whose influence is solely positive; (b) with the second method—a 'general factor' method—we may look for common factors which will be bipolar and therefore have both positive and negative saturations. Where, as in Thurstone's present research, the selection of tests is to a large extent abruptly discontinuous, the former method is evidently the more appropriate. It was, for example, used by Burt and his co-workers nearly twenty years ago in several studies of educational tests, where the subjects tested fall into obviously discontinuous groups. In these early researches the centroid formula (as it is now called) was employed for the first time; and factors very similar to those now reached by Professor Thurstone were elicited (⁴, Tables XVIII-XXIV).

The categories which Thurstone's tests were selected to represent are described as follows: (i) Abstraction (Tests 4-8); (ii) Verbal (9-16, to which should obviously be added 56-60 from the 'unclassified'); (iii) Space (17-25); (iv) Number (30-35); (v) Numerical reasoning (36-39); (vi) Verbal reasoning (40-42); (vii) Spatial reasoning (43-45); (viii) Rote-learning (46-51); (ix) Unclassified, including spelling, grammar and vocabulary (52-60). Thurstone's own grouping thus shows how discontinuous his categories are. To a large extent they coincide with well-established group factors. There can, therefore, be little question that the group-factor method is the natural procedure.

Accordingly, it seemed eminently desirable to test this view by applying the group factor method to Thurstone's table of correlations. The formula used is a modification of the simple summation formula (*viz.*, ⁵, p. 359, equation iv). After eliminating the general factor the remaining

factors are derived from the smaller submatrices of residuals. The only point of difficulty is to determine in advance the lines of division between the several submatrices or clusters, so as to base the general factor on correlations uninfluenced by the one and same group factor. Where the grouping of tests is itself a subject of investigation, we cannot adopt the categories by which the original selection of tests was made; for this would obviously beg the question at issue. The criterion proposed is the degree of resemblance between the various columns of correlations. To study these resemblances we may either calculate the unadjusted inter-columnar correlation or make graphs of the coefficients and judge the resemblances between the contours (cf. 4, fig. 9). Where the correlations between the correlations are non-linear, the latter seems the more reliable as well as the speedier method.

The saturation coefficients obtained by this method are shown in Table I. The first or general factor is responsible for 31 per cent of the variance. On eliminating its effects, there are six submatrices containing significant positive residuals. The group factors derived from these contribute about 2 to 6 per cent of the total variance only. Thus, the general factor is five times as significant as any other.

Professor Spearman, in a paper read at the recent Reading conference, has maintained that Thurstone's table could be fitted by a two-factor analysis and that this procedure would reveal a single general factor. Thurstone, on the other hand, declares: "We cannot report any general common factor in Spearman's sense in the 56 tests that have been analyzed." This is rather surprising, since, in selecting the tests, "special emphasis was laid on those tests which are used as measures of intelligence." Now his Table III does, as a matter of fact, show a 'general common factor in Spearman's sense', i.e., a column of saturation coefficients, all positive, and larger than those in any other column; and its subsequent disappearance is plainly an inevitable result of his method of rotation: this aims, not only at abolishing negative saturations, but also at maximizing the zeros *in every column*, even where the saturations are large and positive throughout. No general factor could survive such a procedure. An analysis by Burt's procedure appears to reconcile the two conclusions: for, with Spearman, we discover a general factor, accountable for more of the total variance than any other, and with Thurstone we discover a number of group-factors having a clear psychological meaning.

In their general nature the group-factors shown in Table I agree almost entirely with those of Professor Thurstone. They prove, indeed, to be much the same as those noted in the earlier researches of Burt and

TABLE I. FACTOR SATURATIONS BY GROUP-FACTOR METHOD.

Test.	G	V	L	A	S	C	M	R	Z
4	.554	.483	—	—	—	—	—	—	—
5	.662	.525	—	—	—	—	—	—	—
9	.293	.531	—	—	—	—	—	—	—
10	.649	.511	—	—	—	—	—	—	—
11	.669	.492	—	—	—	—	—	—	—
16	.611	.437	—	—	—	—	—	—	—
52	.533	.496	—	—	—	—	—	—	—
56	.497	.404	—	—	—	—	—	—	—
58	.398	.832	—	—	—	—	—	—	—
59	.237	.265	—	—	—	—	—	—	—
60	.741	.465	—	—	—	—	—	—	—
12	.605	—	.351	—	—	—	—	—	—
13	.537	—	.548	—	—	—	—	—	—
15	.437	—	.628	—	—	—	—	—	—
57	.688	—	.351	—	—	—	—	—	—
30	.678	—	—	.448	—	—	—	—	—
31	.302	—	—	.649	—	—	—	—	—
32	.395	—	—	.575	—	—	—	—	—
33	.349	—	—	.743	—	—	—	—	—
34	.461	—	—	.641	—	—	—	—	—
35	.565	—	—	.444	—	—	—	—	—
37	.627	—	—	.313	—	—	—	—	—
38	.483	—	—	.465	—	—	—	—	—
39	.683	—	—	.446	—	—	—	—	—
8	.444	—	—	—	.424	—	—	—	—
17	.389	—	—	—	.589	—	—	—	—
18	.495	—	—	—	.606	—	—	—	—
19	.520	—	—	—	.512	—	—	—	—
20	.340	—	—	—	.750	—	—	—	—
21	.670	—	—	—	.489	—	—	—	—
22	.504	—	—	—	.622	—	—	—	—
23	.510	—	—	—	.497	—	—	—	—
24	.565	—	—	—	.453	—	—	—	—
27	.367	—	—	—	.555	—	—	—	—
28	.575	—	—	—	.382	—	—	—	—
29	.561	—	—	—	.336	—	—	—	—
36	.304	—	—	—	.214	—	—	—	—
45	.696	—	—	—	.325	—	—	—	—
53	.299	—	—	—	.525	—	—	—	—
6	.814	—	—	—	—	.436	—	—	—
7	.684	—	—	—	—	.364	—	—	—
14	.657	—	—	—	—	.427	—	—	—
26	.418	—	—	—	—	.549	—	—	—
51	.309	—	—	—	—	.445	—	—	—
46	.361	—	—	—	—	—	.499	—	—
47	.527	—	—	—	—	—	.569	—	—
48	.420	—	—	—	—	—	.457	—	—
49	.472	—	—	—	—	—	.404	—	—
50	.370	—	—	—	—	—	.495	—	—
40	.688	—	—	—	—	—	—	[.575]	—
42	.653	—	—	—	—	—	—	[.575]	—
54	.409	—	—	—	—	—	—	—	[.520]
55	.707	—	—	—	—	—	—	—	[.520]
25	.584	—	—	—	—	—	—	—	—
41	.824	—	—	—	—	—	—	—	—
43	.868	—	—	—	—	—	—	—	—
44	.772	—	—	—	—	—	—	—	—
Per cent Variance	30.80	5.00	1.65	4.58	6.61	1.74	1.79	[1.16]	[.097]

Factor V—Verbal-Literary.

Factor L—Verbal-Linguistic.

Factor A—Arithmetical.

Factor S—Visuo-Spatial.

Factor G—General Factor of Mental Ability.

Factor C—Classification.

Factor M—Memory.

Factor R—Relational.

Factor Z—Audio-Rhythmic.

his co-workers on London school children: there he found, in addition to the general factor, more or less identifiable with 'intelligence', two verbal, one arithmetical, a manual, and (in tests more purely psychological) a factor for memory and a factor or factors for sensory perception. Here no manual factor is discovered: but that is presumably because in a collection of tests to be given by the group procedure Thurstone was unable to include any tests of manual dexterity or skill; the place of the manual or mechanical factor seems largely taken by the spatial factor. But perhaps the most interesting point of agreement between the present table and the earlier results is the presence of *two* distinguishable verbal factors: this moreover accords, not only with the conclusions drawn in the London work (⁴, p. 59), but also, it would seem, with Thurstone's own conclusion. The only important discrepancy between Thurstone's list and ours is that he distinguishes three types of relational or rational factors, whereas we find hardly any significant evidence for one. The reason is clear. If (as Burt has maintained) intelligence is manifested most fully and most clearly in 'activities involving reasoning, i.e., the use of logical relations' (⁷, p. 12), then Thurstone's 'logical factors' are mainly a special manifestation of our general factor. Thurstone does not refer to Alexander's work (⁸; cf. 9, pp. 365-71): but it may be noted that Alexander, who used a similar method of rotation, also confirmed the existence of a general, a verbal, an arithmetical, and a practical factor, and endeavoured to demonstrate their importance for educational and vocational practice.¹

Perhaps, however, the most interesting result of our analysis is this. By the use of a very simple procedure we are able to demonstrate and calculate much the same factors as are demonstrated and calculated by Thurstone. Thurstone's own analysis depends first on making an elaborate formal analysis by the centroid method and then rotating the

¹ Since the foregoing analysis was undertaken, we have learnt that Professor Holzinger has also made an analysis of Thurstone's data on somewhat similar lines⁹. The volume of *Psychometrika* containing this study was not received by our Department until late in the year; hence our investigation was taken in complete independence of Holzinger's. As has elsewhere been pointed out (⁶, p. 361), Holzinger's new method of bi-factor analysis (not his original method) is in general principle largely identical with Burt's earlier group-factor method. The chief differences are, first, that Holzinger allocates the tests on the basis of what he calls a beta-coefficient, and, secondly, that his method of deducing the general factor saturations would appear to depend on a multiplication formula rather than on a summation formula. Neither in his *Student Manual* nor in his previous Reports of the Spearman-Holzinger Trait Committee does he express his method in terms of an actual formula; but the method as described would appear to imply the use of Burt's equation vi instead of his equation iv (⁶, p. 355). In spite of these slight divergences in procedure, our results appear to be closely similar. In each group, however (except those for arithmetic and memory), our own table shows one or two minor additions and one or two omissions as compared with Holzinger's. It may be added that our method, with 9 factors, accounts for more of the total variance than Holzinger's with 10.

axes thus found by a somewhat prolonged and admittedly precarious graphical procedure. The submatrix method reaches the same results directly with one set of simple calculations. Since we have relied on fewer factors, our figures do not fit the observed correlations quite so well as Thurstone's. But of the residuals remaining from our analysis only 2 out of 1,596 are over 0.3. When, as here, the probable errors are high (± 0.07 according to Thurstone), residuals of this size can have no statistical significance, particularly in so huge a table. If a more complete set of saturations were required, giving a slightly closer fit, it could be obtained by carrying the calculation a stage further according to the method described and illustrated in a previous number of this *Journal* (⁶, p. 55).

To educationists one of the most interesting chapters in the monograph is the last. This deals with the uses of mental 'profiles' based on the factor measurements, and suggests the possibility of picking out those individuals who are marked by exceptionally high or exceptionally low performance in some particular factor and therefore might be said to belong to the 'type' which that factor designates. In particular, it is found that many of the individual profiles show an instructive relation to the vocational interests and wishes of the persons they represent: thus the two youths having profiles with the highest relative scores in the factor of verbal relations (what we have called the verbal literary factor) desire to be teachers; others, who have high scores in the visio-spatial factor, wish to be engineers or geologists.

In conclusion we must express our admiration for the great care and thoroughness which has evidently been expended upon this research. It is, indeed, one of the most valuable educational experiments of its kind hitherto carried out. It provides a mass of figures for those who wish to test alternative methods of analysis; and anyone who wishes to be acquainted with the factorial technique in educational research will find this book a most lucid and instructive introduction.

H. J. EYSENCK.

REFERENCES.

- ¹ THURSTONE, L. L. (1935). *The Vectors of the Mind*.—University of Chicago Press.
- ² — CHESIRE, L., and SAFFIR, M. (1933): *Computing Diagrams for the Tetrachoric Correlation Coefficient*.—University of Chicago Bookstore.
- ³ HARTOG, P., RHODES, E. C., and BURT, C. (1936): *Marks of Examiners*.—London: Macmillan and Co.
- ⁴ BURT, C. (1917): *Distribution and Relations of Educational Abilities*.—London: P. S. King and Sons.
- ⁵ — (1938).—*Journal of Psychology*, VI, 339–375.
- ⁶ — (1939).—*This Journal*, IX, 1, 45–71.
- ⁷ — (1911).—*Journal Exper. Pedagogy*, 1, 93–112.
- ⁸ ALEXANDER, W. P. (1935): *Brit. J. Psych. Mon. Suppl.*, XIX.—London: Cambridge University Press.
- ⁹ HOLZINGER, K. J., and HARMAN, H. H. (1938).—*Psychometrika*, III, i, 45–60.

OUTLINES OF RESEARCHES REPORTED IN THESES PRESENTED FOR HIGHER DEGREES OR DIPLOMAS.

THESE OUTLINES MUST BE SUBMITTED THROUGH THE HEAD OF
THE DEPARTMENT IN WHICH THE RESEARCH WAS CARRIED OUT.

A Study of the Development of the Concept of Death.

*Thesis approved for the M.A. degree in the University of London, 1939, from
University College, London.*

By SYLVIA ANTHONY.

The problem of what death means to children at different stages of their development has been studied through three different methods: (1) home records; (2) a story-completion "test," devised at the Institut J.-J. Rousseau, Geneva, for purposes of general psychological diagnosis; (3) recording of relevant responses in Terman-Merrill Intelligence Testing Scale, and the insertion of the word "dead" in the vocabulary list.

The subjects numbered 123; not all were studied by all methods. The range of age and intelligence was wide (3-11+ years, 30-170 I.Q.), but the majority were selected at random from L.C.C. Schools. Twenty-six cases were seen at Child Guidance Clinics.

From the Story-Completion test it was found that children are very ready to refer to death, when in fantasy mood. This confirmed the findings of the devisers of the test, to whom it had come as a surprise. Of the ninety-eight subjects, 51 per cent made at least one explicit reference to death; the percentage was slightly lower among the clinical cases. A marked feature of many such references was the working of the talion idea; reference to a parent's (fantasied) death would be followed by reference to the child's; fantasies of reparation were also observed. A child very much preoccupied with the recent death of his father avoided fantasy-references to death altogether.

Adults' reactions to children's thinking about death are discussed, from a historical and psycho-analytic point of view.

The insertion of "dead" in the T-M vocabulary produced eighty-five responses which fell into five categories, ranging from A (ignorant) to E ("mature"). The great majority fell into a median category—a disproportion which added to the difficulty of drawing conclusions from the statistical data. The trend of change in the children's conception of death bore out the conclusions of Professor Piaget both as to the general form and as to 7-8 being the pivotal age.

From the home records the stages indicated by the categories are illustrated in intimate detail. The transition from A (ignorance) to B (limited concept) is related to the development of the concept of cause and the discrimination of self. The C (junior) stage corresponds to a stage when the idea, at first often found uncharged with emotion, links up with affective complexes already well developed, mainly those associated with (a) birth

and (b) aggression. In the death-aggression complex animal fantasies come to play an important part. The death-birth complex suggests in some children spontaneous ideas of reincarnation.

Anxiety about death is considered as being either chronic or critical: the chronic type is commonly connected with aggressive impulses, and may be activated at any age; the critical type may occur when the inevitability of death for the (discriminated) self is first faced. Cases are quoted where the mother encouraged the child in negation of this inevitability. It is suggested that this may be psychologically justifiable.

A Study of Children's Activity with Plastic Material and Some Interpretations of Play in Infancy.

A thesis accepted as part-qualification for the degree of M.A. in Education (University of Birmingham, 1938).

By EDNA M. HALL.

PLASTIC material (clay and plasticine) was chosen for this study because of its popularity with children and because it offers particularly wide scope for their imagination.

Records were made of the free activity with this material of two hundred children between the ages of five and eight from three schools, two in industrial towns and one in the country. Fifty of these children were chosen for more intensive observation in groups of four or five, the groups consisting of children from varying home conditions and of (a) average or less than average and of (b) superior intelligence at each of the age levels 5+, 6+, 7+. These children were not abnormal in behaviour or emotional development.

The records show certain characteristics of each age group. The play of the five-year-old children was crude, varied and quickly changing and showed disregard for the orthodox use of the material. The play of the six-year-olds was more coherent. In one group of six-year-old children 40 per cent of the work had a story or central interest, while only 8 per cent of the work in a comparable group of five-year-olds was connected. The older children showed greater powers of concentration and a decided appreciation of the possibilities of their material. The seven-year-old children made little use of their material. They preferred conversation.

Of the eighty-eight types of objects made by the two hundred children, by far the most popular was the aeroplane, which was made sixty-one times. (This was in 1937-1938.) Food, cakes, sweets, fruit, etc., were next (forty-five times). Animals of all kinds were third (forty-one times). People—this includes children, soldiers, and babies—were fourth (thirty-seven times).

The choice of subjects of town and country children demonstrated the influence of their different environments.

Repetition was frequently noticed since most of the children returned at least once to a favourite subject.

The play observed could be divided into several types:

- (1) Representation of environment, e.g., children living near a canal made horses, barges, and bridges.
- (2) Expression of a child's particular interest. A boy of six, for example, much interested in football, made football pitches and players.

- (3) Experiment with material such as using clay as cement for bricks.
- (4) Topical play, e.g., logs for bonfire night, pancakes for Shrove Tuesday, presents at Christmas.
- (5) Dramatic and Exuberant Play, e.g., boys playing at "Indians in their den," shooting with clay guns.
- (6) Phantasy. There were not more than six children whose imaginative play was so divorced from reality as to be termed phantasy.

The children's sense of humour was frequently stirred by crudities and exaggerations in the objects they made; e.g., a cat with excessively long whiskers caused much amusement.

Incipient powers of logical and well-founded criticism were apparent in their remarks about each other's work. One boy said of another's house: "It's flat. When you get in you'll be outside."

There were no signs of sadism—no chopping off of the arms or legs of the figures they had made.

An interpretation of the activities of these children on psycho-analytic lines would emphasize the significance of the latent content of their play. But if this play is seen in its natural setting of the surroundings of the children and the events of their life, its relation with this environment is so marked and so constant as to be worthy of special consideration. Further, a child's careful perfection of detail, supposedly symbolic of restitution, is frequently found to be the result of real interest and purposeful observation. The free activities of these children showed their concern with reality, and some wish-fulfilment as well as a good deal of sheer fun. The "manifest content" has its own significance, and the need for other analysis is rarely apparent.

AN EDUCATIONAL FAILURE : A REPLY.

To the Editor.

By F. H. HAYWARD.

DEAR SIR,

In response to a review of my book, *An Educational Failure*, in the June number of your journal, may I say that the following paragraphs give my theoretical bases for which the reviewer asked :

- (a) There exists a mass or crowd emotion, whether due to Herd Instinct, Imitation, Telepathy, the call of the Great Being, or (if you will) the inspiration of the Holy Ghost.
- (b) The factors of novelty and the employment of specially qualified teachers play their part in this high temperature result, as also do :
 - (i) the more careful preparation demanded and forthcoming for assembly work over the more private classroom work ;
 - (ii) the general high quality of the pieces chosen for recitation, etc., and the fact that they are linked together by a dominating idea ;
 - (iii) the dignity of the themes (great persons, ideas and events) ;
 - (iv) the response of the artists and performers to the call of the mass emotion aforesaid ; the greater ability on their part to "let themselves go" passionately in a large audience rather than in a small.

Yours faithfully,

F. H. HAYWARD.

THE STUDY OF EVACUATED CHILDREN : A SUGGESTED ENQUIRY.

THE évacuation of so many children gives a great opportunity to teachers to make useful observations on the personality or temperament of the children, especially in nursery or other schools where teachers are actually living with the children in the same building. Some teachers may have time to record exact observations, particularly on the changes brought about by évacuation, or on the revelation of innate tendencies or individual differences, brought out by the new communal life. If such observations are sent to the Editor of this journal it might be possible to make a useful collective report, especially if quite independent and spontaneous reports afford similar evidence.

Useful guidance in the making of observations and records will be found in *The Educational Guidance of the School Child* (Evans Bros., London), especially in Appendix I, or, for Nursery School Children in *The Social and Emotional Development of the Pre-School Child* by Katherine M. Bridges (Kegan Paul, London).

Suggestions for detailed observations and records will be given in an article in the next number of this journal.

BOOK REVIEWS.

American Psychology before William James: By JAY WHARTON FAY.
(Rutgers University Press, 1939, pp. 240. \$2.50.)

This modest but readable and informative volume fills a real gap in the historical treatment of psychology. The book covers the period from 1640 to 1890 (the date of James' *Principles*). It is divided into three parts, dealing respectively with "The Period of Theology and Moral Philosophy (1640-1776)"; "The Period of Intellectual Philosophy" (1776-1861), and "The Period of British and German Influence" (1861-1890). In the third period the parting of the ways can perhaps be said to occur in 1887, when Stanley Hall in the memorable first number of the *American Journal of Psychology* reviewed "the ultimate broadside of the expiring philosophical discipline."

Of the many interesting, but to-day little known, authors to which this history refers, we may mention a few of special importance. Samuel Johnson in 1752 produced the very important *Elementa Philosophica*, which is amazingly wide in scope, and in some respects in advance of contemporary European works. Of particular interest to readers of this journal is his attitude towards educational problems and the treatment of children generally. "It seems evident that those little creatures, from the beginning, do consider, reflect and think a prodigious deal more than we are commonly apt to imagine . . . We ought to think little children to be persons of much more importance than we usually apprehend them to be; and how indulgent we should be to their inquisitive curiosity as being strangers; and with how much candour, patience and care, ought we to bear with them and instruct them."

Thomas C. Upham (1799-1872) anticipated—of course, under other names—many features of modern psychology, including 'mental set,' 'individual differences,' 'introversion and extraversion,' 'rationalization,' 'displacement,' or 'the emergence of suppressed desires in perverted forms,' and the 'James-Lange theory of emotions.'

Finally, we may mention that remarkable pair, the two Rushes, the father Benjamin Rush (1745-1813) and the son James Rush (1786-1869), almost comparable in brilliance to the two Mills on the other side of the Atlantic. Benjamin Rush used the technique of suggestion a century before the Nancy school. The younger Rush, according to Dr. Fay, may be regarded as the first behaviourist. His very formidable *magnum opus* was only published in 1865 (the date is wrongly given on p. 140) under the resounding title of "Brief outline of an Analysis of the Human Intellect, intended to rectify the Scholastic and Vulgar Perversions of the National Purpose and the Method of Thinking; by rejecting altogether the Theoretic Confusion, the Unmeaning Arrangement, and Indefinite Nomenclature of the Metaphysician." As it ran to 930 pages, it is perhaps fortunate that he confined himself to 'a brief outline.'

These few examples must be sufficient to indicate the contents of this fresh and invigorating book, which forms a distinct contribution to psychological scholarship. It contains 47 pages of literary and biographical notes and references, a useful chronological table of American and relevant European works, a bibliography of 'primary sources' and a good index of names (though not of subjects). Dr. Fay's volume undoubtedly fulfils the aim which is indicated in the quotation from Pliny the younger that figures on its title page; "I hold it a noble task to rescue from oblivion those who deserve to be eternally remembered."

J.C.F.

Introduction to Educational Psychology: By W. C. TROW. (The Riverside Press; George Allen and Unwin, Ltd., pp. 417. 8s. 6d. net.)

This book deserves to gain a wide circulation. It is written in a clear and fluent style. Only the lists of references appended to the chapters indicate its American origin. It is a well balanced exposition and thoroughly integrated selection of substantiated observations and alternative explanatory theories. The book itself is

an excellent illustration of the author's own chapter on the scientific attitude. The order of presentation is unusual but symptomatic of the more recent balance of emphasis. For example, the opening account of the living creature from the point of view of its needs and strivings is followed by a consideration of emotion and the autonomic nervous system; the C.N.S. is described 150 pages later. Each chapter has a good summary. There is a glossary of scientific terms and a good index. It is an attractive publication to which full justice cannot be done in a short notice.

E.J.G.B.

General Psychology: By GILLILAND, MORGAN AND STEVENS. (Harrap. 8s. 6d.)

This is a fairly comprehensive text-book for the student beginning a course on psychology. After an introductory chapter on the nature and purpose of psychology, the authors go on to explain the physiological and sensory bases of behaviour, and, adhering closely throughout to this physiological background, they deal with the processes of attention, learning, thinking, and emotional behaviour. A brief account of intelligence tests is given and there are chapters on Personality, Social Behaviour and Mental Health. A short bibliography is given at the close of each chapter, consisting almost exclusively of the works of American psychologists, and the book appears to be designed primarily for the American student. There is little reference to the Gestalt point of view in the chapters on learning and memory and almost as little to the work of Freud in the short account that is given of dreams. Nevertheless as an elementary text-book the book is clear and concise, and should be helpful to the young student.

M.C.P.

The Psychology of Social Movements: By PRYNS HOPKINS. (London: Geo. Allen and Unwin, pp. 284, 10s. 6d. net.)

Dr. Pryns Hopkins has written an interesting book to which he has given the sub-title of "A Psycho-Analytic View of Society." After introductory chapters in which he summarizes Freud's theory of psychology as an analysis of human nature, he proceeds to apply this teaching to many social questions, in chapters on "Man and Cruelty," "Man and Violence," "Man and his Bread," "Man and his Possessions," "Man and his Children," "Man and his Sensory Enjoyment," "Man and his Knowledge." He summarizes the position in a final chapter, "The Road to Happiness." The author has certainly succeeded in writing a book which will give rise to much discussion and criticism, with resultant modification and clarification.

Psychological Methods of Healing: By WILLIAM BROWN. (London: University of London Press, pp. vi+224. 7s. 6d. net.)

This volume is a general introduction to psychotherapy. The author has included brief accounts of the thought and practice of some of the leading schools, but he also gives the basis of his own system. Dr. Brown's comments on the theories of leading psychotherapists are always interesting, and students will turn with attention to those chapters dealing with hypnosis, suggestion, relaxation and mental analysis. Some chapters are devoted to the consideration of some applications of principles of psychotherapy to social, political and even international problems.

A valuable addition to the book is a selected bibliography of one hundred books to give assistance for further reading. This list is in itself instructive.

Manual of Psychology: By G. F. STOUT, revised by C. A. MACE. Fifth Edition. (University Tutorial Press, pp. 708. 12s. 6d.)

In times when so many flimsy books are written on psychology it is encouraging to the teacher of psychology to find so solid, penetrating, and comprehensive an exposition as is given by Professor Stout's Manual, continuing to be reprinted and revised after so many years.

The present edition is enhanced in value by an appendix by R. H. Thouless, in which is given an excellent exposition of the Gestalt psychology, with some discussion of its relation to Professor Stout's own teaching. Old students of Stout have rarely been able to get very excited about the Gestalt psychology, and will be prepared for a number of the anticipations of that school which are in Stout's work, and are here indicated.

Finally, the edition includes a most welcome supplementary note by Stout himself, in which he discusses some paragraphs of Thouless's Appendix, and expounds the relationship of his own teaching to that of the Gestalt psychology.

Fear and Depression: By a Medical Psychologist. (Cornish Bros., pp. 90. 5s.)

In this booklet, which discusses the causes of fear and depression and the methods of self-treatment, the author regards resentment towards others as a main source of mental trouble. He gives a large number of psychological suggestions and ideas for self-control, and many hints for the carrying out of a kind of auto-suggestion. Most of these seem to be fairly sound and not to depend too much upon the speculative, and often unreliable, psychology that is scattered among the more theoretical parts.

As so often happens in books of this kind, what may be occasional coincidences are treated as serious evidence.

Psychological Factors in Marital Happiness: By LEWIS M. TERMAN AND OTHERS. (McGraw-Hill Publishing Co., 1938, pp. xii+473. 24s.)

It was a bold thing to attempt the investigation of such a delicate problem, but Terman approached the matter in a manner which could hardly be improved upon. He has managed to collect anonymous replies to his questionnaire from over one thousand married couples (including the preliminary enquiry) and over one hundred divorced couples.

The danger of selection, which would result in the replies not being typical, was avoided to some extent by calling various kinds of groups together and getting the questionnaires filled in by the whole group. Though the groups were of different types we cannot however be quite sure that these were completely typical. As Terman himself points out, these societies which joined in helping with the questionnaires contained on the whole a rather large proportion of persons interested in "uplift" activities, or in matters of self-improvement, Church Groups, Mothers' Clubs, etc. Also they were of middle or upper middle class, and better educated than the average.

A very large and varied questionnaire was submitted and the frequency and importance of particular answers calculated. The correlations of marital happiness as estimated by the persons themselves were also calculated.

The report abounds in interesting results. As we might expect, when it comes to making up quarrels, most mates think that they themselves "give in" rather than their opposites, for example, 16 per cent of the wives assert that they "give in" whereas only 4 per cent of the husbands acknowledge wifely submission. On the whole the estimates of happiness of the marriages are fairly cheerful ones; for example, out of about 800 husbands 232 described their marriage as "exceedingly happy," and 292 as "decidedly above the average." The number of women who describe the marriage as exceedingly happy was appreciably greater—274. Of course, from these marriages the unsuccessful ones which had resulted in broken homes or divorce would be excluded.

As to the causes of unhappiness, it proved that the frequency of grievances was not as important as their seriousness. The correlation of only 0.6 between the "happiness scores" of the husbands and those of the wives shows that the happiness of the one spouse is to a surprising degree independent of the happiness of the other. Childless women past middle age show a slight tendency to be less happy than the average, but childless men of this age tend to have happiness scores above the average. No doubt this latter result is due to the fact that the chief interest of many wives is deflected from husband to children when the latter arrive.

The correlation of income with happiness scores was zero: but it must be remembered that the majority of the couples would be comfortably off.

Of slight importance to marital happiness were such factors as adequate sex instruction, length of pre-marital acquaintance, and early marriage.

"As to relative mental ability, the most favourable situation is equality or near equality. Marked mental superiority of husband makes for happiness in the wife but for unhappiness in the husband; marked inferiority in the husband makes the wife unhappy but does not greatly affect the husband."

Happiness of the two mates in childhood and happiness in their parents are important "background circumstances" where the marriages were happy. But surely these would be highly correlated with innately happy temperaments, which would tend to result in the married life being happy.

The influence of sex factors on marital happiness seemed to be less than is often asserted, though the husband's happiness is *negatively* correlated with the wife's "prudishness or excessive modesty."

On the whole, while there are several gaps we should like to have seen filled, the investigation is a very notable contribution to this important psychological and sociological question.

C.W.V.

Measurement: By WILLIAM A. MCCALL. (The Macmillan Company, New York, 1939, pp. xiv+535. 18s. net.)

This book is a revision of the author's *How to Measure in Education* which appeared in 1922. A comparison of the present volume with the original shows the nature and extent of recent developments in educational measurements in the United States. It also shows the extent and variety of the tests which are daily administered in American schools and the degree of complexity of the testing movement in the United States. Many of the tests described are but little known in this country.

Professor McCall has played a conspicuous part in these developments. His style of writing is occasionally racy and forceful. This is decidedly helpful, as, from the nature of his task, many of the pages are packed with dry-as-dust, yet necessary, rules of the technique of testing. Some of these pages would prove boring to the experienced tester, yet they have perforce to be included, otherwise the book would be incomplete for the student.

In spite of the prolific appearance of books on testing in recent years, the book can be strongly recommended even to those who possess the original edition, for, as already indicated, extensive additions have been made. In justice to the would-be purchaser, however, the reviewer feels bounds to impart some information. In the first place, although the title "Measurement" has brevity to commend it, yet the book is very largely concerned with educational measurement in the United States. Secondly, although the science of measurement is supposed to be international, yet there is no mention in the index of any British writer on measurement, if we except a solitary reference to Pearson's product moment formula, and another to the Spearman-Brown prophecy formula! Thirdly, the title "Measurement" inevitably invites comparison with "Measurement in Psychology," which was the title of Mr. R. J. Bartlett's presidential address to the Psychology Section of the British Association this year. Such a comparison again shows that Professor McCall's field is that of educational measurement. Every chapter has been carefully prepared and the author throughout is an enthusiastic champion of the possibilities of measurement. Thus in reference to the many factors which tend to decrease the reliability of a test he writes: "By the proper juggling of these factors a pupil's ability in many tests may be practically returned to his original ability" (p. 57).

Very instructive and even amusing is the example on p. 313 when two of the author's close friends rated his personality traits. Very refreshing is the way in which he upholds subjective measurement. "Nor can we condemn the measurement as being inaccurate simply because two raters disagree. They may disagree widely and yet both be perfectly correct. We should not expect perfect agreement, since the two persons doing the rating do not themselves have identical personalities. The logic of ordinary methods of determining reliability does not quite hold" (p. 311).

The comprehensive nature of the author's testing plans may be seen by a brief mention of the chief divisions, namely, (1) Intelligence Test; (2) Educational Background Test; (3) The Comprehensive Achievement Test; (4) The School Practices Questionnaire. Some of these tests will be novel to British readers. The volume is divided into eight books:

(1) *Place of Measurement in Education.* Here the author's philosophy of measurement is unfolded. Suffice it to say here that the object of testing is to achieve happiness. Happiness as a criterion is here discussed.

(2) *Criteria for the Selection and Construction of Tests.*

(3) *Use of Standard Tests for Grouping Pupils.*

(4) *Programme of Measurement for Progressive Schools.*

(5) *Guidance and Evaluation of Teaching by Measurement.*

(6) *School Marks and Reports.*

(7) *Presentation of Test Results.*

(8) *How to Scale Tests and Compute Statistical Measures.*

LL.W.J.

Political Arithmetic: A Symposium of Population Studies: Edited by LANCELOT HOGBEN. (George Allen and Unwin, pp. 531. 30s.)

The volume opens with a fascinating introduction by the editor. Here is mapped out the programme of inquiries and researches carried out by a team of experts from the Department of Social Biology, University of London. Some of the results have already appeared in journals which have only a limited circulation among specialists, but about half of the material is presented for the first time. As would be expected, Professor Hogben's style is incisive and his criticisms of certain writers on economics are lively. Interesting is the account of the founders of the study of population. The names of the three founders of British demography—Graunt, Petty, and Halley—appear in the original list of Fellows of the Royal Society when it received its charter in 1662. What kind of man was William Petty, the author of *Political Arithmetick*, which appeared in 1691? "By turns cabin boy, hawker of sham jewellery, seaman, inventor, physician, Fellow and Vice-Principal of Brasenose, Professor of Anatomy at Oxford, and of Music at Gresham College, surveyor, Member of Parliament, landed proprietor, philosopher, statistician, and political economist!"

Professor Hogben's masterly treatment should be supplemented by reading Chapter VII, "British Demographers' Opinions on Fertility, 1660 to 1760," by Dr. Kuczynski.

As the second part of the volume will probably receive more attention from readers of this *Journal* a bare mention of the first six chapters must suffice:

- I. The Internal Decline of Fertility, by Dr. Kuczynski.
- II. The effect of Present Trends in Fertility and Mortality upon the Future Population of Great Britain and upon its Age Composition, by Dr. Enid Charles.
- III. Differential Fertility in England and Wales during the Past Two Decades, by Dr. Enid Charles and Miss Pearl Moshinsky.
- IV. Changes in Fertility in England and Wales, 1851 to 1931, by Mr. D. V. Glass.
- V. The Changing Structure of the Family in Australia, by Dr. Charles.
- VI. Marriage Frequency and Economic Fluctuations in England and Wales, 1851 to 1934, by Mr. Glass.

Turning to the second part, Chapter VIII, Ability and Opportunity in English Education, by Mr. J. L. Gray and Miss Pearl Moshinsky is of immediate interest to students of psychology. The Otis Group Advanced Test, Form A, was administered to more than ten thousand children in London schools of widely varied educational and social type. The results are sufficiently disquieting to demand further investigations, e.g., "taking children of equally high ability, seven fee-paying pupils will receive a higher education for every one free pupil. Conversely, if we consider children who fall below the selected level of ability, for every one free pupil who is afforded the opportunity of a higher education, there are one hundred and sixty-two fee-paying pupils who enjoy the same advantages."

It may be doubted, however, how far the Otis test can be accepted as a standard for this type of research with English children. The results obtained seem to be at variance with the widely known results obtained by Professor Burt, and also by others who have made a comparative study of children from different social strata. In a study where political considerations are apt to intrude, it is important to secure a valid test which should be administered by experts under strict conditions.

In Chapter IX the same two authors apply their findings with the Otis test in relation to Parental Occupation.

Chapter X, by Mr. Glass and Mr. Gray, is entitled Opportunity and the Older Universities. Although the investigators here did not have access to as many sources as they could have wished, yet there is no reason to doubt that the facts here tabulated are substantially correct. For, unlike the Otis results, they are facts which follow immediately from the original data. Thus, in 1933-34, ex-public elementary schoolboys contributed only 12 per cent of Oxford and 13 per cent of Cambridge undergraduates. Public and private schools obtained 78 per cent of all college entrance scholarships at Oxford. About 32 per cent of entrance scholarships at the older universities are awarded in the subject of classics. Of these, 92 per cent go to pupils of public schools. Public schools have one full-time teacher to every fourteen boys and county secondary schools one to every twenty-three. Seventy-five per cent of the full-time teaching staff of public schools are graduates of Oxford and Cambridge.

It is well to have these facts tabulated even if they could have been generally surmised. Generations of county school boys have been under no delusions as to the formidable handicaps confronting them when they elected to compete for such scholarships. It is essential then to have access to such tables as have been so ably presented in this chapter. The next step is more difficult, namely, to agree on the reforms which are warranted by these facts. Occasionally statements are made which are not strictly true. "At present the only way in which an ex-elementary schoolboy can proceed to the university is by the award of a scholarship or by obtaining some similar form of financial aid from some outside source" (p. 422). This obviously requires qualification as some parents of ex-elementary schoolboys were actually able to afford help.

Chapter XI, The Relation between Initial and Maximum Earnings and Differential Fertility of Skilled and Unskilled Wage-Earners, by Dr. Charles and Mr. D. Morgan, may be strongly commended to industrial psychologists.

Chapter XII, The Distribution of the Blood Groups and its bearing on the Concept of Race, by Mr. Allison Davis, is a technical account of the data concerning the four main blood-groups which deserves the study of all geneticists and anthropologists.

LL.W.J.

The Year Book of Education, 1939. General Editor: HARLEY V. USILL.
(Evans Bros., Ltd., pp. 823. 35s.)

This book came into our hands just too late for review in the June number, but it is a type of year book, as no doubt many of our readers already know, which is by no means out of date the year after, or a good many years after, the year of publication. For it contains, as usual, not only a section dealing with current statistics in the British Commonwealth of Nations, Europe, and the U.S.A., and a survey of finances in this country, but a series of sections on important present-day problems, treated in such a way as to be of permanent interest.

The present volume includes a large section giving a survey of senior and central schools, prepared under the direction of Professor H. R. Hamley. Another very important section is that which deals with technical education. A very wide survey of this is given, general problems first being treated by writers of different nationalities and then a number of chapters giving reports of technical education in Germany, Italy, and several other European countries, in the Dominions and the United States. Finally there is a section on current research in education edited by Professor Cavanagh, these first contributions in this part of the *Year Book* being chiefly of an historical type. The most interesting, however, of all the sections to most readers of this journal will be that headed "The Philosophies of Education," which includes a chapter by Lord Stamp on "Education and the Survival of Democracy," one on "Education in Eire," by Professor Corcoran, and, most interesting of all at

the present time, the chapter on "The Conditions and Content of the New Order of German Education," by Dr. Gräfe, the Director of the German Academic Exchange Service, Berlin. This article, written, no doubt, late in 1938, gives a succinct and carefully reasoned account of the basis of the educational schemes under the Nazi Government. It emphasizes the revolt from the over-intellectualized education in Germany and the greater attention to the training of character, which is spoken of as a remarkable change, but, of course, is very familiar to us in Great Britain. This training of character in Germany, however, is to be largely a training for readiness to give "unconditional sacrifice under the new banner of the National Socialist Liberation Movement." The object of physical training is not merely that we shall have "bodies sound to the core" but a training of the "will power and ability to arrive at decisions," a phrasing which suggests still the dominance of the doctrine of formal training in this aspect of education, although Dr. Gräfe explicitly discards it in an earlier section when dealing with intellectual education. The exposition of the main ideas underlying the prescribing of a "land year" for the youth of great urban centres is brief but compact. Equally illuminating is the section on the purpose of the secondary school reorganization into one type of school (apart from a few Gymnasias) and the elimination of the social bias in selection based upon the parents' financial ability, an idea, of course, with which we have long been familiar and putting into practice to a large extent in our ordinary secondary schools.

A notable sentence is this: "All those pupils who suffer from incurable physical diseases or other defects, as well as those who continuously offend propriety, *esprit de corps*, discipline and uprightness, or whose intelligence does not reach a certain minimum limit, are excluded from secondary schools." Notable also is the sentence in the final paragraph: "The new organization of the German school system is determined by a new, national political conception of man, which introduces the deciding phase of our era in German national and educational history. The next period of history and its bearers will be determined in a wide measure by this conception of man." This section on German education alone would make this an important volume of this *Year Book* series.

The Nature of Creative Activity: By VIKTOR LÖWENFELD. Translated by O. A. Oeser. (Kegan Paul, pp. xvii+272. 21s.)

This book seems to provide the most noteworthy contribution to the body of facts on which to base a theory of aesthetics since Kerschensztein published his study of children's drawings in 1905. Professor Löwenfeld has brought out quite clearly the distinction between two types of creative activity: one dependent on visual experience, the other on tactual and kinæsthetic experience. His method consists in comparing the artistic efforts of the blind with those of the sighted—especially the early artistic efforts. The blind, of course, can deal only with sculpture; they can neither draw nor paint. But the abnormally weak-sighted—those who are so weak-sighted that they can only see objects an inch or two away from the eye and can rarely see an object, not even a flat picture, as a whole—can both draw and paint after a fashion; and the point to note is that their productions exhibit just the same characteristics, and just the same departures from photographic faithfulness, as do the productions of the blind.

So in fact do the drawings of young children, of savages, of the primitives, and of the ultra-modern school of artists. In all these we find a challenge to the popular conception of artistic merit. For the man in the street judges a work of art by a standard which is purely visual. It looks like some original model, or it does not; if it does it is good, if it does not it is bad.

Professor Löwenfeld disposes of this theory out of hand; not so much by argument as by the presentation of abundant data which demonstrate that "haptic" experience, as he calls it, supplies another criterion which is probably more fundamental. By haptic experience he means somatic sensations other than visual, together with a consciousness of the self. Haptic art has a scale of values quite independent of visual standards, and is rarely free from symbolism of some kind. It plays havoc with our notions of time and space, and is always deeply imbued with the personality of the artist.

The author's clear exposition of these non-visual factors in a work of art does not merely throw light on certain extreme schools of expressionism which flourish to-day, but reveals some of the more recondite merits of the old masters. In fact, the book is of considerable value both to those who study art and to those who practise it.

P.B.B.

The Assistant Master Speaks: By C. D'O. GOWAN and Others. (Kegan Paul, pp. 291. 7s. 6d.)

This volume, consisting of essays by twelve assistant masters in a variety of secondary schools, forms the latest of a series published by Kegan Paul of which the previous publications were *The Head Master Speaks* and *The Head Mistress Speaks*.

The assistant master occupies a position which, although perhaps more sheltered from direct criticisms than his Head, is more closely in touch with the boys themselves and his views are likely to throw a new light on present-day educational problems. It is good to find that the writers of these essays are not wholly drawn from the larger public schools, but that the day grammar and secondary schools are well represented. The range of topics is wide.

In matters such as reforms in the curriculum and the new outlook on discipline, it is evident that the assistant master has an important contribution to make, but the three most characteristic contributions and those which only assistant masters could have written are *Ourselves as Others see us*, *Beast or God*, and *The Menace of Paganism*.

The first deals with the status of the assistant master in society and the harmful effects of segregation. The second examines the factors which affect the effectiveness of a master particularly as he grows older. It abounds in provocative statements of which the following is an example: "The most effective part of an assistant master's life seems to me to be short and early. I think that few over thirty teach as well and honestly as those between twenty-six and that age."

The last of these three deals with the failure real or apparent of school Christianity. "I believe that Christianity in schools as in the world has been rejected because it has not been fairly tried . . . it seems to me an inescapable conclusion that the blame for this failure must be laid mainly at the door of the assistant master." These are hard words, but the writer makes out a strong case to support his contention.

To judge from the contents of this book, assistant masters are by no means complacent.

Educational Traditions in the English-speaking Countries: By N. HANS, D.Lit. (Evans Bros. 2s. 6d. net.)

This extract from the *Year Book of Education* for 1936 would have been greatly improved if the publishers had added an index.

Whether or not we think that Dr. Hans has made his case, we shall agree that in the thesis before us he has shown immense knowledge of his subject, and that while many writers have pointed to differences between educational systems, he has gone further by putting stress on the origins of the systems and upon what appears to link them together. He claims that five factors contribute to national unity, and unity may be effected where any three are operative at once. Race and language act directly for union, but religion is complex; although in Scotland and in South Africa it unifies, it has caused division in Ireland and in Canada. Territorial distribution and political independence complete the list of factors, but they are less potent than the first named, certainly when these three are combined.

We cannot deal with all the material presented by this encyclopædic essay, we must bow before its mass of detailed information. One or two printers' errors should be corrected. We may doubt if some facts that are stated are really proofs of the conclusions arrived at, but Dr. Hans may be forgiven these claims having regard to the immense labour he has put in to sorting and weighing a mass of material dealing with English-speaking countries throughout the world.

One should not omit to mention the useful bibliography which has been given.

A.P.B.

Education for Citizenship in Elementary Schools: with a Foreword by EARL BALDWIN OF BEWDLEY, K.G., and KENNETH W. M. LINDSAY, M.P. Issued by the Association for Education in Citizenship. (Oxford University Press, 1939, pp. 312. 4s. 6d.)

The Association for Education in Citizenship defines education for citizenship as: "Training in the moral qualities necessary for citizens of a democracy, the encouragement of clear thinking in everyday affairs and the acquisition of a knowledge of the modern world."

In 1935 the Association published a volume, *Education for Citizenship in Secondary Schools*, in which consideration was given as to how these aims could be carried out with pupils of secondary and public school age. The book under review considers the same task with regard to children under the age of fifteen.

The book is divided into three sections. The first three chapters present the aims and theory of education. This is followed by the main body of the volume consisting of sixteen chapters, which deal with methods and which set out clearly and concisely what is being done, and in some cases what might in addition be done in the schools. The final section—the appendix—contains details of schemes of work which are being successfully used in schools and which illustrate admirably the many points raised in the foregoing chapters.

Together, the three sections contributed by either administrators or by successful and experienced teachers constitute a whole which is logical, stimulating and suggestive and which supplies a want which has long been outstanding.

G.N.G.

Education of the Handicapped: Editors—MERLE E. FRAMPTON and HUGH GRANT ROWELL. (Sixteen Contributors.) Vol. I. (Harrap and Co., London, 1939, pp. 260, 7s. 6d. net.)

The object of the editors and contributors to this work has been to devote the first volume to a chronological review of the gradual recognition of the specific problems involved in the care and education of the handicapped, while the second will deal with the social philosophy which underlies the treatment of these problems. The scheme is thus an exceedingly interesting one, and one which, judging by the volume now issued, is likely to be of very great interest to all members of the community upon whom the responsibility of dealing with the problem must ultimately rest. Recognition of the fact that the handicapped is a citizen and that as such he has a right to every privilege which his citizenship connotes despite his disability, has been strangely slow, as also has its corollary that it is to the interest of the community that every member thereof shall be as physically and mentally fit as the knowledge of the day can make him.

As the editors point out, the movement towards this end has always been the result of the self-dedication and enthusiasm of some great and far-seeing individual, and the main object of Vol. I is therefore to a large extent to give an account of the contributions of such enthusiasts towards a larger understanding of the problems involved, with such biographical details as can be given in a short summary.

The term "handicapped" has a wide connotation, and includes not only the crippled, those with serious defects of sight and hearing, and the mentally inefficient, but also the stammerer and those whose handicap arises from heart disease, tuberculosis and malnutrition. There is also a section which, under the term "the socially handicapped," deals with the general question of juvenile delinquency and the prevention of crime.

A composite volume by a number of contributors always produces a certain amount of discrepancy as regards statements of facts, dates and the like—a difficulty felt by the editors, who remark in the preface: "Much basic evidence is obscure or inaccessible . . . in some instances given dates and the spelling of certain terms and names are at variance. Versions of the same incident are frequently misleading." This is true, and one could, if critically inclined, point to occasional inaccuracies, but such inaccuracies amidst such a wealth of material detract but little from the excellence and usefulness of this book, which should be in the hands of every teacher, whether of the handicapped or the child who is sound and fit. We shall look forward to the appearance of the second volume.

G.A.A.

Bias and Education for Democracy: By MICHAEL STEWART. (Oxford University Press, 1938, pp. 48. 1s.)

The most frequent objection made to education for citizenship or education for democracy is that politics should, at all cost, be avoided in the school, it being generally felt that, since there is a lacking in the subjects bearing on citizenship of a quantitative accuracy common to mathematics and the physical sciences, the affairs of the contemporary world cannot be taught without bias.

Owing largely to the activities of the totalitarian States, "propaganda" has become a much misused word and it is encouraging to read Mr. Michael Stewart's closely reasoned analysis of its full meaning.

If democracy is to survive the fervent challenge offered to it by totalitarianism, it is surely necessary that the virtues of a democracy should be strongly put before the nation's young. The liberty of the individual in speech, thought and action is surely of such infinite value that it warrants the use of "a type of propaganda which seeks support for a given view by marshalling reasoned argument in its support. . . . The spreading in this manner of belief in liberty, justice, truth and kindliness, the standards of conduct on which the future civilization now depends, is justifiable and specially necessary at the present time." (p. 12).

The many difficulties in the way of such teaching together with the obvious dangers of excessive bias are fully discussed. At the same time, it is appreciated that these dangers are not so great as it would at first appear, for the teaching of public affairs to young children will consist largely of the presentation of facts, and the older pupils will have a livelier critical faculty and will, at the same time, come into contact with a diversity of opinion from outside the school as well as from within.

Mr. Stewart is firm in his belief that open discussion on controversial matters is essential to education for citizenship. The many difficulties and dangers which accompany such discussions must be faced up to and overcome.

Bias and Education for Democracy should be of interest and value to every teacher or student of education whether he be actively engaged in the teaching of subjects allied to citizenship or not, for it reveals the possibilities and values of free discussion in all its aspects.

G.N.G.

Moses and Monotheism: By SIGMUND FREUD. Translated from the German by Katherine Jones. (Hogarth Press, pp. 223. 8s. 6d.)

In this book Dr. Freud attempts to find support in history for his theory that the origin of religion is to be found in the Oedipus Complex. The historical premises on which he builds his argument are highly disputable. They are that Moses was an Egyptian, that the monotheism he gave to the Hebrews was that of the heretic Pharaoh Ikhnaton, that he was murdered by the people of his adoption, and that the submerged religion came to the surface centuries later in the teaching of the Prophets. The fact that Moses was murdered was first "discovered" some twenty years ago by Sellin, in an attempt to prove that Moses was the "Servant of the Lord" of *Isaiah* liii. Dr. Freud is probably Sellin's first disciple, and in the meantime Sellin himself has so far abandoned his theory as that he no longer believes that Moses was the "Servant of the Lord," which was the only reason that led him to discover that Moses was murdered. The psychologist may make of this book what he will; it is unlikely that it will carry conviction to the historian.

C.R.N.

The Young Teacher's Handbook: By GEORGE H. HOLROYD. (Elkin, Mathews and Marrot. 2s. 6d.)

This is a concise and practical guide written primarily for the student just leaving college. It deals clearly and comprehensively with most of the branches of the teacher's work, including games and out-of-school activities; various professional organizations are described and an appendix on the Burnham scale of salaries is added. The needs of the male teacher are considered more fully than those of the women, but the handbook should be useful to all those on the threshold of a teaching career.

M.C.P.

War and Democracy. Essays on the Causes and Prevention of War : By E. F. M. DURBIN, JOHN BOWLBY, and Others. (London : Kegan Paul, pp. viii+360. 10s. 6d.)

This group of essays consists of six essays, *War and its Causes, 1815-1914*, by Ivor Thomas ; *Nationalism and Capitalism*, by D. P. T. Jay ; *Pacifism and Collective Security*, by R. B. Fraser ; *Collective Security and Dictatorship*, by R. H. S. Crossman ; *Pacifism, Marxism and Peace*, by G. Catlin ; and *Personal Aggressiveness and War*, by E. F. M. Durbin and John Bowlby.

The last-mentioned essay appears first in the volume and is the one which is of most interest probably to readers of this journal. The essay itself is followed by a long appendix giving the evidence on which the conclusions in the essay are based. The evidence is largely based on the work of Zuckermann, Susan Isaacs, and the psycho-analysts, with many references to anthropology. The essay and appendix are extremely suggestive.

It is interesting to note that this essay and appendix have been reissued in a separate volume from the same publishing house under the title *Personal Aggressiveness and War* at 5s.

The Generations : By EMMANUEL MILLER. (London : Faber and Faber, pp. 276, 7s. 6d. net.)

Dr. Miller, well-known for his psychological work and publications, has here essayed a different task. He is concerned with the many problems connected with the relationship between different generations, particularly those between parents and children. The subject-matter is based on a series of lectures on "The Problems of Parenthood." Some of the topics discussed are "The Background of Marriage," "The Emotional and Intellectual Development of the Child," "Puberty and Adolescence," and "The Future of the Family."

On these and many other contingent subjects the author writes in a very stimulating way and formulates many new problems. His attempted solutions are not always satisfying and some of his generalizations seem to be based on insufficient evidence.

Studies in Arithmetic. Reports on Investigations relating to Present Practice and Teaching Methods in the Primary School. Volume I. Publication of the Scottish Council for Research in Education. (University of London Press, pp. lxix+144. 5s. net.)

The researches reported in this volume are a genuine attempt to discover certain facts which, if known, would affect the teaching of arithmetic.

The first part deals with the present practice in a representative sample of Scottish schools—such questions as time allocation and methods used in teaching the various topics. The analysis of the questionnaire is valuable ; perhaps the final very useful summary would be improved if in comparing the numbers favouring the several methods, percentages had been used.

The second part of the book deals with an enquiry into the relative difficulty of basic number facts—this is a careful piece of work and much needed. Teachers are apt to give their pupils indiscriminate drill in number combinations, yet if, as it appears, $7+9$ is more difficult to memorize than $7+7$, more drill should be given to the former combination than the latter.

These researches are needed and should find place in the mathematics section of a junior school teachers' library and on the shelves of every training college.

M.H.

Education for Citizenship : By G. A. ROWAN ROBINSON. (Association for Education in Citizenship, London, 1939, pp. 16. 6d.)

This small pamphlet is a lecture in printed form. It should be useful to those possessing a slight acquaintance with the school system of Germany, though it is too brief, perhaps, for the absolutely uninitiated. The author's interpretation is reasonably fair.

A.T.

Scripture Teaching To-day: By M. VIVIAN HUGHES. (Student Christian Movement Press, 1939, pp. 253. 5s.)

This is a practical book, written by one who has had an unusually wide experience in the teaching of Scripture, and in observing others teach it, and who wishes "to pass on some of the good things witnessed, and to refer to the bad only for contrast."

The style of the book is curiously reminiscent of that inimitable book, *Swiss Family Robinson*, beloved by the children of a more leisurely century. Indeed, the author is, in the realm of Scripture teaching, as omniscient, as wise, and as encouraging to the tribe of earnest teachers as was papa in every branch of knowledge to his large and questioning family!

The author has real insight into the specific difficulties that arise in teaching this difficult subject. The suggestions for solution are of interest and worth. A particularly valuable element is the many references to literature helpful for the elucidation and illustration of Bible passages.

Teachers of Scripture should find the book most helpful.

M.H.

Arithmetic in Action: By E. BRIDEOAKE and I. D. GROVES. (University of London Press, Ltd., pp. 128. 4s. 6d.)

Those interested in practical arithmetic will find this book useful. Miss Brideoake and Miss Groves outline a complete scheme of arithmetic from the reception to the transition stages; an arithmetic, however, bereft of fear and confusion, transformed to a series of exercises productive of self-confidence and real comprehension of the numerical processes. The scheme embraces all that can reasonably be required in an infants' school syllabus. Knowledge of the fundamentals is consolidated by games and puzzles.

A short chapter describes the results obtained by this method; results not easily measured or tabulated, since the scheme includes no examination papers, but evidenced by the child's confident use of his tools, his comprehension of principles and his power to attack problems. The chapter headed "Hints on Transition" enhances the value of the book for those who wish to test the practical method but fear confusion during the transition.

A First Course in Statistics: By Prof. E. F. LINDQUIST. (pp. xi+226. 8s. 6d.)

A Study Manual for a First Course in Statistics: By Prof. E. F. LINDQUIST. (pp. 120. 5s.) (London: Geo. G. Harrap and Co., Ltd.)

Professor Lindquist, of the State University of Iowa, has written an extremely useful course in statistics particularly designed to show their use in education and psychology. These two books may be used as "a method of teaching" a first course in statistics. The usual topics are discussed simply and clearly, and the necessary theory is elucidated. It is important to note also that the author emphasizes the limitations of statistical techniques. This is particularly essential at the present time.

Superstition and Society: By R. MONEY-KYRLE. (pp. x+161. 4s. 6d. net.)

Civilisation, War and Death. Selections from three works by SIGMUND FREUD. (pp. viii+102. 3s. 6d. net.) (London: The Hogarth Press and the Institute of Psycho-Analysis.)

These two volumes are Nos. 3 and 4 in the series of Psycho-Analytical Epitomes. The former deals, in an interesting way, with many of the psycho-analytic contributions to anthropology under the headings of mythology, totemism and exogamy, animism and magic, the origin and development of culture, and culture and education.

The latter contains selections from three of Freud's works, *Thoughts for the Times on War and Death*, *Civilisation and its Discontents*, and *Why War?*

Both volumes will be welcomed by students and general readers.

War and Peace: By WILLIAM BROWN. (London: A. and C. Black, 1939, pp. xvi+93. 5s. net.)

This series of essays in psychological analysis represents the results of Dr. Brown's very careful study over many years of the factors making for war and of the means whereby the outbreak of war may be averted. The subject is divided into War and Psychology; International Relations; Psychology, Ethics and Group Morality; Unconscious Forces and War; and Psychological Conditions of Peace. Each topic is discussed carefully and many interesting and important psychological aspects are illuminated as a result of the author's investigations. A very useful book, as it raises so many subjects for further examination.

Scholars or Soldiers? Aims and Results of "Nazi" Education: By THEODOR WILHELM. (Duncker and Humblot, Berlin, 1939, pp. 102.)

This useful little book is a reprint from the *International Education Review*, Vol. VIII, 1939, No. 2. It is the work of one actively engaged in German education with a complete grasp of his subject and the further advantage of an understanding of his readers. Having already described the educational ladder in Germany in some detail he now attempts the much more difficult task of interpreting the philosophy of the new system. This is one of the few reliable publications on the subject.

A.T.

Journey to a War: By W. H. AUDEN and CHRISTOPHER ISHERWOOD. (Faber and Faber, pp. 301. 12s. 6d. net.)

This is an unorthodox travel book, but loses nothing in value by its intimate style. It gives a vivid picture of China in her struggle against Japanese imperialism and of the many types which make up her population.

OTHER PUBLICATIONS RECEIVED.

- Histoires Varides*: Isabel Hawkes. (Macmillan and Co. 1s. 9d.)
Four French Plays: N. W. H. Scott. (Macmillan and Co. 2s. 6d.)
Selection from Modern Poets: M. Woolman. (Macmillan and Co. 1s. 9d.)
The Boundaries of Science: John Macmurray. (Faber and Faber Ltd. 7s. 6d.)
Children from Seed to Saplings: Martha M. Reynolds. (McGraw-Hill Publishing Co. 15s.)
Prose Selections with Exercises (first year): H. Alderton Pink. (Macmillan and Co. 1s. 6d.)
Prose Selection with Exercises (second year): M. Alderton Pink. (Macmillan and Co. 1s. 9d.)
Latin Prose Composition: H. E. Gould and J. Whiteley. (Macmillan and Co. 2s. 6d.)
More Poems Old and New: selected by A. Cairncross and J. Scobbie. (Macmillan and Co. 2s. 6d.)
Short Modern Plays: selected by S. R. Littlewood. (Macmillan and Co. 2s. 6d.)
Higher School Certificate Latin Prose: A. H. Nash-Williams. (Macmillan and Co. 3s.)
Poetry and Appreciation: A. F. Scott. (Macmillan and Co. 3s. 6d.)
The Tempest: Edited by J. R. Sutherland. (Oxford University Press. 2s.)
Language Games for Infants and Juniors: B. E. Houghton. (Macmillan and Co. 2s. 6d.)
A Midsummer Night's Dream: Edited by F. C. Horwood. (Oxford University Press. 2s.)
Diversions for the Arithmetic Class: E. H. Renwick. (Macmillan and Co. 2s.)
Sixty Letter and Reading Games: Jane
British Association of Mathematical Teachers: by
 W. F. Sheppard. (Cambridge Un
The Heating, Ventilation and Lighting of School Buildings: by W. P. Seymour. (Oxford University Press. 12s. 6d.)

